

視差情報を用いたグラフカットによる 4次元光線空間の領域分割

三原 基^{1,a)} 船富 卓哉¹ 田中 賢一郎^{2,1} 久保 尋之¹ 長原 一³ 向川 康博¹

1. はじめに

画像処理やコンピュータビジョンの分野において、画像から注目する物体の領域を抽出する技術は領域分割と呼ばれる。2次元画像を対象とした領域分割手法としては、Interactive Graph Cuts[1] が有効であることが知られており、様々な画像処理ソフトウェアにも実装されている。近年、シーンの新しい記録手法として4次元光線空間のイメージング技術が発展しているが、これに対する領域分割としては、同手法を適用することが有望であると考えられる。現状では、4次元光線空間を2次元画像の集合として表現し、これらへ複数回繰り返すことが考えられるが、ユーザにとっては大きな手間となるだけでなく、光線空間全体に一貫した領域分割結果が得られるとも限らない。そこで本稿では、Interactive Graph Cuts を4次元光線空間へと拡張した領域分割手法を提案し、その有効性を示す。

2. 2次元画像の領域分割

本章では、領域分割手法の1つである Interactive Graph Cuts について説明する。 $x-y$ 平面上の2次元画像 P における、あるピクセル $p = (x, y)$ のラベルを $l_{x,y} \in \{1, \dots, l\}$ とする。Interactive Graph Cuts は、 P の一部に入力として与えられるラベル i の正解ラベルを持つピクセルの集合 S_i ($i = 1, \dots, l$) から、画像 P に含まれるすべてのピクセルに対して、最適なラベル集合 $L = \{l_{x,y}\} \forall x, y \in P$ を求める手法である。互いに隣接する2つのピクセル (p, q) の集合を N としたとき、エネルギー関数 $E(L)$ を次式で表し、それを最小化する L を求める。

$$E(L) = \sum_{p \in P} R_p(l_p) + \lambda \sum_{\{p,q\} \in N} B_{p,q}(l_p, l_q) \quad (1)$$

$$R_p(i) = -\ln \Pr(I_p | S_i) \quad (2)$$

$$B_{p,q}(l_p, l_q) = \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \delta(l_p, l_q) \quad (3)$$

なお、 λ は第1項と第2項の比率を表すパラメータであ

¹ 奈良先端科学技術大学院大学

² 大阪大学

³ 九州大学

a) mihara.hajime.lt2@is.naist.jp

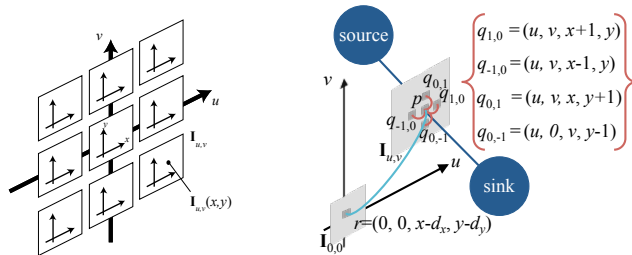


図 1 4次元光線空間

図 2 光線空間のグラフ表現

る。ここで、 I_p, I_q はピクセル p, q での画像の輝度、 $l_p, l_q \in \{1, \dots, i, \dots, l\}$ は p, q が持つラベル、 σ は定数である。また、 $\delta(l_p, l_q)$ は $l_p \neq l_q$ のとき 1、それ以外を 0 とする関数である。式 (2) は、輝度値に関する事後確率を S_i から計算し、 p がラベル i を持つ尤度を I_p によって評価するデータ項、式 (3) は、隣接ピクセル同士が似ていた場合と同じラベルを割り当てやすくする平滑化項である。Interactive Graph Cuts では、式 (1) を表現するグラフを作成し、min-cut/max-flow algorithm[2] を適用することによって、式 (1) を最小化する L を決定する。

3. 提案手法

3.1 問題設定

本研究での領域分割の対象とする光線空間は、図 1 に示すような、4次元光線空間 I を視点 (u, v) から観測した2次元画像 $I_{u,v}$ の集合として表現する。本手法では、4次元光線空間 I に加え、視点 (u, v) における画像 $I_{u,v}$ の各ピクセルに対して基準視点 $(0, 0)$ における画像 $I_{0,0}$ との間で定義される視差ベクトル画像 $D_{u,v}$ が既知であるものとし、ユーザは基準視点における画像 $I_{0,0}$ の一部に正解ラベル S_i を与えるものとする。

3.2 エネルギー関数

2次元画像に対する Interactive Graph Cuts では、 $x-y$ 方向の隣接関係を考慮したエネルギー関数が定義された。本研究では4次元光線空間を一貫して領域分割するために、 $x-y$ 方向に加え $u-v$ 方向にも拡張したエネルギー関数を定義する。

光線空間 I 上に割り当てられるラベルの集合を $L =$

$\{l_{u,v,x,y}\} \forall u, v, x, y \in \mathbf{I}$ とし, $l_{u,v,x,y} \in \{1, \dots, l\}$ である. ここで, 4次元光線空間上の点 p を (u, v, x, y) とするとき, このピクセルの近傍を (x, y) 方向と (u, v) 方向で分けて考える. まず, (x, y) 方向の近傍として, p と画像 $\mathbf{I}_{u,v}$ 内の p の隣接ピクセルの集合を \mathbf{N} とする. また, (u, v) 方向については, p と \mathbf{D}_p によって得られる $\mathbf{I}_{0,0}$ 上の p の対応点 r との対の集合を \mathbf{C} とする. 対応点 $r \in \mathbf{C}$ は $\mathbf{D}_p = (d_x, d_y)$ のとき, $(0, 0, x - d_x, y - d_y)$ によって求められる. ただし, ほとんどの場合に p と r は同じ領域に対応づくが, p が領域の境界付近だった場合, 隠れの影響により r が異なる領域に対応づく場合がある. このとき, 式 (1) を拡張すれば, グラフカットで最小化するエネルギー関数を, 以下のように定義できる.

$$E(\mathbf{L}) = \sum_{p \in \mathbf{I}} R_p(l_p) + \lambda_1 \sum_{\{p,q\} \in \mathbf{N}} B_{p,q}(l_p, l_q) + \lambda_2 \sum_{\{p,r\} \in \mathbf{C}} B_{p,r}(l_p, l_q) \quad (4)$$

なお, λ_1, λ_2 は第 1 項に対する第 2, 3 項の比重とする. ここで, 式 (4) の第 1 項, 第 2 項はそれぞれ式 (2), 式 (3) と同様の関数と定義する. ここで新たに追加した第 3 項は, 視差ベクトルによって求められた p の対応点 r との平滑化項を表している. p と r が同じ領域に対応づいた場合は, I_p と I_q が同じ色になると考えられるため, それらを同じラベルに割り当てやすくするが, 違う領域に対応づいた場合, p と r の色は異なると考えられるため, それらを違うラベルに割り当てやすくする効果を持つ.

3.3 グラフ表現とエネルギー関数の最小化

式 (4) で定義したエネルギー関数を表現するため, 光線空間全体から図 2 のようなグラフを構成し, min-cut/max-flow algorithm を用いてこれを分割することによって, 式 (4) の最小化を行う. グラフには, 4次元光線空間の各ピクセルに対応した頂点と, ラベル $i = 1, \dots, l$ に対応するターミナル T_i が存在する. このとき, 各枝の重みは, 表 1 に示す値と設定する. なお, 表中の $i, j (i \neq j)$ はラベル, K は十分大きな定数を表す. このようにして得られたグラフに対し min-cut/max-flow algorithm を適用することで関数を最小化するラベル割り当て \mathbf{L} を求めることができ, 4次元光線空間の領域分割が可能となる.

表 1 各枝に与える重み

edge	weight	for
$\{p, q\}$	$B_{p,q}(l_p, l_q)$	$\{p, q\} \in \mathbf{N}$
$\{p, r\}$	$B_{p,r}(l_p, l_q)$	$\{p, r\} \in \mathbf{C}$
$\{p, T_i\}$	$\lambda_2 \cdot R_p(l_p = i)$	$p \notin \cup \mathbf{S}_i$
	K	$p \in \mathbf{S}_i$
	0	$p \in \mathbf{S}_j$

4. 実験

本手法の有効性を示すために実験を行った. 実験で用い

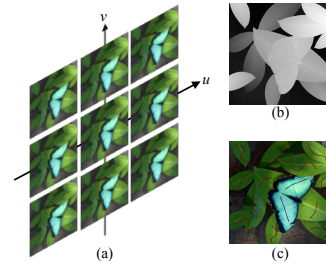


図 3 入力 (a): \mathbf{I} , (b): $\mathbf{D}_{u,v}$, (c): 青色が \mathbf{S}_1 , 赤色が \mathbf{S}_2

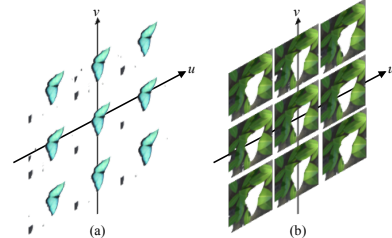


図 4 領域分割結果 (a): ラベル 1, (b): ラベル 2

る 4次元光線空間および視差情報は [3] で提供されている 4次元光線空間データセットを用いた. なお, 4次元光線空間の視点数は $3 \times 3 = 9$ とした. 図 3 に示す 4次元光線空間, ユーザのブラシストロークによる正解ラベル (注目物体を表す \mathbf{S}_1 , それ以外を表す \mathbf{S}_2 の 2 種類), 視差情報を入力とし, 尤度 P_r を YCrCb 色空間のうち Cr と Cb の色分布確率で表すものとする. 4次元光線空間の領域分割を行った結果を図 4 に示す. 結果より, 基準視点画像 $\mathbf{I}_{0,0}$ の一部だけに正解ラベルを与えるだけで, 周辺視点 (u, v) における画像に対しても一貫した領域分割を達成した. これより, 提案手法が光線空間全体の領域分割に有効であることがわかる.

5. おわりに

本稿では, グラフカットによる領域分割を 4次元光線空間に適用する手法を提案し, 実験を行った. その結果, 4次元光線空間全体の領域分割を実現した. 今後はさらなる高精度化のため, 輝度以外の情報を特徴とした手法の開発を行い, 本手法との比較を行う予定である.

謝辞

本研究は科学研究費補助金 (基盤 A, No. 25240027) の助成を受けたものである.

参考文献

- [1] Yuri Boykov, and Marie-Pierre Jolly, "Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images," in *Proc. International Conf. on Computer Vision*, Vol. 1, pp.105–112, 2001.
- [2] Yuri Boykov and Vladimir Kolmogorov, "An Experimental Comparison of Min-cut/Max-flow Algorithms for Energy Minimization in Vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 9, pp. 1124–1137, 2004.
- [3] S. Wanner, S. Meister, B. Goldluecke, "Datasets and Benchmarks for Densely Sampled 4D Light Fields," in *Vision, Modelling and Visualization*, pp.225–226, 2013.