

# 顔認識アルゴリズムを用いた擬似立体視の実現

## マルコフ確率場を用いた地面の深度推定

高松 真†

† 東京電機大学 〒120-8551 東京都足立区千住旭町

E-mail: †16ec068@ms. dendai. ac. jp

キーワード 擬似立体視, 相対深度, ハフ変換, マルコフ確率場, Belief Propagation, ステレオマッチング

Makoto TAKAMATSU†

† Faculty of Engineering, First University Yamada 1-2-3, Minato-ku, Tokyo, 105-0123 Japan

E-mail: †16ec068@ms. dendai. ac. jp

### 1. ま え が き

単眼画像の立体視は困難な問題である。一枚の静止画像から立体視を実現する問題を解決する手段として、ハードウェアで実際の環境の深度を測定し、測定結果を単眼画像に加える方法 [1, 2] や深層学習が注目されている [3, 4, 5]。我々は、ハードウェアを使用せず、ソフトウェアのみで単眼画像の擬似立体視を実現するために、物体認識アルゴリズムを用いて被写体を認識、被写体のサイズから相対深度を推定し、各被写体の深度を等間隔に割り当て擬似立体視を実現する方法について検討した [6, 7]。

地表を映した写真では、地面の深度がすべての領域で等しく、また各被写体の相対深度が異なるため、擬似立体視の見え方に違和感を感じることがあった。そこで、筆者は3点解決策を提案する。1つ目は、空は無限遠点に存在する領域であり、相対深度で扱うことは不可能な領域である。したがって、無限遠点である空の領域と相対深度として扱える被写体領域を区別するための消失線推定を行うことを提案する。2つ目は、地面の相対深度を消失線から画像の先頭まで、被写体の人数で地面の深度を等分することで擬似立体視の見え方への違和感を減らすという目的で提案する。3つ目は、正確な地面の深度は不明であるため、左右に移動させた画像でマルコフ確率場によるステレオマッチングを行い、深度が画一的でない地面を作成することを提案する。このとき、被写体の相対深度は決まっているため、地面のみ画一化されていない深度を求めることを意図している。

### 2. 我々の従来法

#### 2.1 顔 認 識

顔認識アルゴリズムである Biola-Jones アルゴリズムを用いて1枚の人物集合写真から顔を認識、検出した顔のサイズを正方形のバインディングボックスで囲む。図は顔のバインディングボックスで囲んである図である。

#### 2.2 相対深度算出

被写体  $i$  の深度を以下の式:

$$D_i = \frac{255S_0}{S_i}$$

で表す。ただし、 $S_0$  は最も後ろの被写体の顔サイズを、 $S_i$  は任意の被写体の顔サイズを表す。

#### 2.3 擬似立体視作成

各被写体の深度を使い、被写体を左右に移動させステレオビジョンを作成するために、移動量を算出する。移動量を以下の式:

$$P_i = M(1 - \frac{D_i}{255})$$

で表す。作成したステレオビジョンにアナグリフ処理を行う。図 ww は作成したアナグリフ画像である。

### 3. 現在行っていること/ 提案

#### 3.1 消失線分析

今までの私たちの研究では、最も後ろの被写体の深度を背景に適用していた。したがって、実際には大きく距離が離れている背景と被写体を同じ深度で擬似立体視を実現した。現在、筆者は最も後ろの被写体と背景を明確に分離するため、ハフ変換

を用いて消失線を検出した。ハフ変換とは以下の式

$$\hat{\rho} = x \cos \hat{\theta} + y \sin \hat{\theta} \quad (1)$$

で表される。 $\hat{\rho}$  は、原点から直線までの符号付き距離、 $\hat{\theta}$  は原点から直線への垂線である。

式 (1) は  $xy$  画像空間中の点を  $\rho \theta$  パラメータ空間に写像すると点の数だけ直線を描くことができ、これらの直線は  $\rho \theta$  パラメータ空間の 1 点  $(\hat{\rho}, \hat{\theta})$  で交差することを示し、この交差する 1 点の座標を検出すれば  $xy$  画像空間中の直線を検出できる。 $\rho \theta$  パラメータ空間中の平行線は画像面上では一点で交わる。この 1 点を消失点と呼ぶ [8]。

図 X の黒線はハフ変換によって得られた消失線である。

### 3.2 地面における深度グラデーション問題解決への提案 1

ハフ変換を用いて消失線を引いた。この消失線以上の領域を深度 255 とし、深度 255 から前景にかけて深度を線形にかけ、被写体の深度を被写体の靴先から靴先までを等分し深度をかける。図 q に等分に地面に深度をかけた図である。

### 3.3 地面における深度グラデーション問題解決への提案 2

我々はいままでスーパーピクセルを用いて、被写体を切り抜き深度マップを作成してきた。スーパーピクセルを用いることは色や位置に類似性を持つピクセルを結合しスーパーピクセルとすることで計算時間を短縮することができるという利点がある。ただし、筆者は被写体を画像から切り抜き、被写体を前景、切り抜かれた画像を背景とする二値化的な擬似立体視をより立体的な画像とするためにマルコフ確率場を用いることを提案する。

マルコフ確率場は画像工学ではグラフカットや画像復元、領域分割などに使われる手法である。

$$E(X) = \sum_{v \in V} g_v(X_v) + \sum_{u, v \in E} h_{uv}(X_u, X_v) \quad (2)$$

式 (3) はピクセルの並びをグラフとしてとらえ、そのグラフのエネルギ関数を最小化するための手法である (詳しい説明は別の報告書を参照)。ただし、マルコフ確率場は NP 困難なアルゴリズムであるので、近似アルゴリズムである Belief Propagation を用いた。

図 Y はマルコフ確率場を使って作成したステレオビジョンの深度マップである。本研究の意義である、単眼カメラでステレオビジョンを実現するという観点から、筆者は相対深度にもとづきピクセルを移動させた左右画像をマルコフ確率場に入力し、深度マップを作成することを提案する。なぜなら、マルコフ確率場を使用した深度マップは、単眼カメラからの深度推定は画像の環境および画像の全体構成に関する事前知識に基づいている。マルコフ確率場は、似ている外観のものは、密集して配置された場合に同様の深度分布を持つと仮定するため、周辺分布を計算できる [9, 10, 11]。という利点があるからである。

スーパーピクセルによって被写体を切り抜き、被写体を左右に移動させ作成したステレオビジョンをマルコフ確率場によって地面の深度グラデーションを調べるために深度推定をおこな

う。すでに被写体の深度を固定した状態で環境の深度マップを作成する。

## 文 献

- [1]
- [2] 画像情報教育振興協会:デジタル画像処理, pp. 238-243(2015)
- [3]
- [4]
- [5]
- [6]
- [7]
- [8]
- [9] A. Saxena and J. Schulte, Andrew Y. Ng, "Depth Estimation using Monocular and Stereo Cues, " IJCAI. 2007.
- [10] F. Liu and C. Shen, G. Lin, "Deep Convolutional Neural Fields for Depth Estimation from a Single Image, " CVPR 2015.
- [11]
- [12]
- [13]