

# 顔認識アルゴリズムを用いた擬似立体視の実現

## 被写体の立体感をいかに作るか？

高松 真<sup>†</sup>

<sup>†</sup> 東京電機大学 〒120-8551 東京都足立区千住旭町  
E-mail: <sup>†</sup>16ec068@ms.dendai.ac.jp

キーワード 擬似立体視, 相対深度, ハフ変換, 被写界深度, ステレオマッチング

Makoto TAKAMATSU<sup>†</sup>

<sup>†</sup> Faculty of Engineering, First University Asahicho 5, Adachi-ku, Tokyo, 120-8551 Japan  
E-mail: <sup>†</sup>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] や深層学習が注目されている [2,3]。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、被写体の顔サイズから相対深度を推定し、各被写体の深度を割り当て擬似立体視を実現する方法について検討した [ ]。

我々の今までの研究では、地面の深度がすべての領域で等しく、反対に、各被写体の相対深度が異なるため、作成した擬似立体視を見ると被写体が地面から浮かんで見えるなどの違和感を感じることがあった。そこで本報告書で、立体的な見え方を実現するために筆者らは3点解決策を提案する。

1 つ目は、ハフ変換を用いて被写界深度の範囲を定量的に測定する。つまり、被写界深度の範囲とそれを超える部分との間を消失点検出を用いて区別する。被写界深度とは、ピントを合わせた部分の前後のピントがあっているように見える範囲のことである。

2 つ目は、画像の下端から消失点まで地面の深度を変化させる。これは擬似立体視の見え方への違和感を減らすためである。その際、地面の深度グラデーションのかけ方について、線形と非線形で比較を行う。(現在、地面への深度グラデーションは非線形で行う方法はプログラムが完成している)

3 つ目は、アナグリフ画像を作成した際に見られる黒い縁取り部分をなくす作業である。黒い部分を無くしてより自然な立

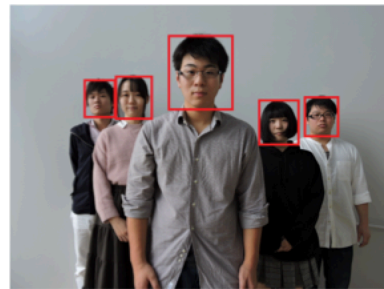


図 1 顔 認 識

体感を生み出すため Blure texture という手法を行う。Blure texture は、動く物体に Blure(ボケ) のかかったテクスチャをマッピングする技術である。シルエットが鮮明でテクスチャのつなぎ目が目立つときに使われる手法である [8][9][10][11]。

本報告書の構成を次に説明する。第 2 章が我々の従来法、第 3 章が消失点検出 (ハフ変換)、第 4 章は地面の深度グラデーション、第 5 章は Blure Texture について説明する。

### 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四角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:

$$D_i = \frac{255S_0}{S_i} \quad (1)$$



図 2 深度画像



図 3 ステレオビジョン作成  
??



図 4 アナグリフ画像

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:

$$P_i = M(1 - \frac{D_i}{255}) \quad (2)$$

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 ?? は作成したアナグリフ画像である.

## 3. 消失点検出

今までの私たちの研究では, 最も後ろの被写体深度を 255 とし, 最も後ろの被写体と背景に深度値 255 を適用した. したがって, 実際には大きく距離が離れている背景と最も後ろの被写体を同じ深度で扱い擬似立体視を実現した. 現在は, 被写界深度の上部とそれを越える部分との境界をハフ変換を用いて消失線検出する. ハフ変換とは以下の式

$$\hat{\rho} = x \cos \hat{\theta} + y \sin \hat{\theta} \quad (3)$$

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

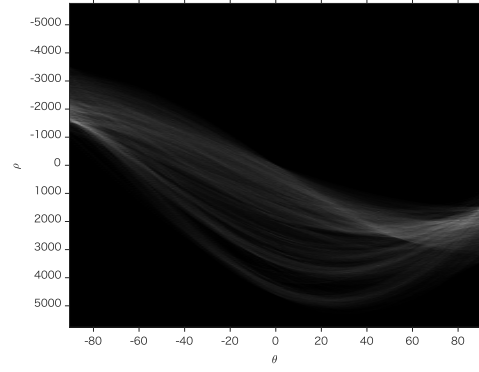


図 5



図 6 消失線検出

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. 画像空間中の直線を検出している様子が図 ?? である.

$\rho \theta$  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 [4].

図 ?? の赤線はハフ変換によって得られた消失線である.

## 4. 地面領域の深度グラデーション

### 4.1 線形バージョン

ハフ変換を用いて消失点を検出した. この消失点より上の領域を深度 255 とし, 画像下端から消失点にかけて深度を線形にかける.

### 4.2 非線形バージョン

線形バージョンと同様にハフ変換を用いて, 消失点を検出する. 画像下端から消失点まで深度グラデーションをかける. 消失点より上は 255 である. 画像は消失点に近くなるにつれて急峻になることを踏まえて画像の深度を非線形でかける.  $\log x$  は収束するまでの距離が長いので, 消失点に近くなると現れる急峻さを表現できない. そこで,  $\log x$  よりも収束速度がはやい  $\sqrt{x}$  を用いる.

## 5. Blure texture

### 文 献

- [1] Julien N.P.Martel , "An active Approach to Solving the Stereo Matching Problem using Event-Based Sensors." 2018.
- [2] E.Shelaharmer,J.Long and T.Darrel"Fully Convolutional Networks for Semantic Segmentation."CVPR 2015.
- [3] Y.Luo "Single View Stereo Matching."2018
- [4] 画像情報教育振興協会:デジタル画像処理,pp.238-243(2015)
- [5] A.Saxena and J.Schulte,Andrew Y.Ng , "Depth Estimation using Monocular and Stereo Cues," IJCAI.2007.
- [6] J.Sun,H.Shum ,and N.Zheng,"Stereo Matching Using Belief Propagation."
- [7] Z.Liu,X.Li ,and P.Luo,"Deep Learning Markov Random Field for Semantic Segmentation."2017.
- [8] Z.Wu,"Deep Markov Random Field for Image Modeling."2015
- [9] Morgan McGuire,and Padraic Hennessy,Michael Bukowski,Brian Osman, "A Reconstruction Filter for Plausible Motion Blur." i3D(2012)
- [10] Jean-Philippe Guertin,and Morgan McGuire,Derek Nowrouzezahrai, "A Fast and Stable Feature-Aware Motion Blur Filter." HPG '14 Proceedings of High Performance Graphics(2014)
- [11] Xuejiao Luo ,and Nestor Z. Salamon, Elmar Eisemann , "Adding Motion Blur to Still Images." Graphics Interface Conferenhce (2018)