

# 物体認識システムを用いた被写体サイズ推定による疑似立体視の検討

高松 真<sup>†</sup> 中村 文香<sup>†</sup>  
<sup>†</sup> 東京電機大学工学部情報通信工学科

岩田 英三郎<sup>††</sup> 長谷川 誠<sup>†</sup>  
<sup>††</sup> ユニバーサルロボット株式会社

## 1. はじめに

立体映像機器の普及により立体視映像を撮影・表示する需要が高まっている。この立体視映像は、一般に、三次元計測機器を用いて撮影される。しかし、これらの機器は十分に普及していない。一方、既に撮影されている複数の静止画像から立体視映像・画像を生成する方法も種々検討されており、一般的なデジタルカメラでの撮影が可能である。しかし、過去に撮影された静止画像一枚の立体視はできない。一枚の静止画像から立体視を実現する方法として、著者らは、静止画像に写っている複数の被写体を領域分割して各領域の凸包を算出し、凸包と領域との重なりを考察することによって、被写体の前後関係を推定、立体視する方法を提案している[1]。しかし、被写体間の前後関係は判定できるが、深度は人手で指定する必要がある。従来法は立体の高精度な再構築ができるが、著者らの方法では簡易な立体視が可能となる。

ここでは、物体認識システムを用いて被写体を認識、被写体のサイズから深度を推定し、立体視を実現する方法について検討する。顔認識システムを用いて、人物の集合写真から顔を認識し、顔のサイズから、各人物の深度を推定する。深度に応じて、人物を左右に移動させた右眼用・左眼用画像を生成し、アナグリフで立体視を実現する。これまで撮影された膨大な画像や映像、単眼のデジタルカメラでこれから撮影される画像や映像の疑似立体視が可能となる。

## 2. 顔認識と立体視画像生成

画像中から顔を認識、顔のサイズを算出する。図1(a)に原画像、図1(b)に顔認識の結果を示す。ここでは、顔を正方形バインディングボックスで示し、顔のサイズは正方形一片の長さとする。人物 $i$ の深度 $D_i$ を、

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

とする。ただし、ここでの $S_0$ は最背人物の顔サイズを、 $S_i$ は人物 $i$ の顔サイズを示す。深度 $D_i$ は0から255までの値をとり、最背人物の深度は255となる。なお、顔の実寸サイズは個人ごとに異なるが、ここでは等しいと仮定して深度を算出する。図1(c)に示すように領域分割し、深度画像を図1(d)に示す。人物 $i$ の左右に動かす移動量 $p_i$ は、



図1. (a)原画像, (b)顔認識, (c)領域分割, (d)深度画像, (e)左眼画像, (f)右眼画像



図2. アナグリフによる立体視

$$P_i = M \left( 1 - \frac{D_i}{255} \right) \quad [\text{pixel}] \quad (2)$$

である。ただし、ここでの $M$ は深度0における移動量[ $\text{pixel}$ ]を示す。人物を左右に移動させて、図1(e)と(f)に示すように右眼用・左眼用画像を生成する。なお、右眼用画像の場合は左へ移動させ、左眼用画像の場合は右へ移動させる。図2に示すアナグリフ画像を生成し、立体視を実現する。また、立体視はWiggle stereoscopyで表示することも可能である。

## 参考文献

[1] 樋口栄作, 岩田英三郎, 長谷川誠, “被写体の前後推定による単一画像からの立体視検討,” 映像情報メディア学会冬季大会, 24B-1, Dec. 2016.