

# 学 位 論 文 の 要 旨

論文題目 Detection of facial feature points in three-dimensional space  
for meal support equipment

(和訳) 食事支援装置のための3次元空間における顔特徴点の検出

氏 名 彭 博 印

In this doctoral thesis, we propose a method for detecting facial feature points in three-dimensional space on the premise of application to meal support equipment for the purpose of supporting disabled persons and elderly people.

Chapter 1 describes the background of the research, existing research, and the significance of this research. Currently, in Japan, the declining birthrate and the aging of society are advancing, the shortage of carers are concerned, and solving the problem of the shortage of workers is one important issue. Especially, assistance of meal, which is one of the pleasure of care receivers, depends largely on the caregiver. In the place such as nursing homes, it always happens that a small number of carers are assisting a large number of meal support. From such a social background, the development of a meal support manipulator has been carried out up to now. However, the main role of the meal support manipulator is to carry food to the mouth, and the users need to move their face in order to put the food into the oral cavity by themselves. Therefore, there is a possibility of giving a burden to the users. In this paper, we propose a method to detect the position of facial feature points in three dimensional space and aim to apply it to meal support manipulator.

In chapter 2, we described the outline of the developed system and the detection of the three-dimensional position by the stereo camera. In this system, two cameras were arranged side by side, and three-dimensional position measurement of the point of object was performed by using parallax from the obtained two images. As a result of measuring the angle of view and optical axis, distance between two cameras, and further calibrating of the image distortion, it showed that we could measure the position of the target with an error within  $\pm 1\%$  in the three dimensional space.

In chapter 3, we described the method of automatic detection of face, nostrils and mouth and the results of verification experiments. In order to extract the face of a person from the obtained image, the skin color information in the RGB color space was examined for features under different bright environments. As a result, regardless of the brightness, the component of R was found to be stronger in skin, nostrils and lips than in other areas, and the difference between R and B components was very large. According to these features, formulation was made focusing on the relationship between values, difference, and ratio of RGB, and we found the suitable conditions for

skin color. On the other hand, when detecting a face by image processing, it is important to recognize the whole face as one connected component, but it is often not recognized as a connected component depending on the state of the face such as a shadow or eyeglasses. Therefore, rough information of the original image was used, and mosaic processing was applied to the image. For an obtained image of  $640 \times 480$  pixels,  $16 \times 16$  pixels were translated into one block and obtained image was divided into  $40 \times 30$  blocks, and the representative color of each block was obtained from the average color of the representative points. Blocks whose representative colors satisfy the condition of the skin color were extracted and smoothed further to estimate the rough face area. In the estimated face region, binarization and labeling processing were performed on the original image, the nostril candidates were narrowed down and nostrils were detected from the aspect ratio and area of each connected component. On the other hand, note that there were many dark areas around the eyes, mouth, and jaws, but the area around the nostrils has a high proportion of bright skin, and the area around the nostrils is estimated by finding the proportion of skin color in the face area. This characteristic was used to confirm the detected nostril position, and it could prevent misdetection of nostril positions. From the obtained nostril position, we estimated the region around the mouth and constructed a system that can recognize the opening and closing state of the mouth by binarization processing. As a result of the verification experiment by the subject, it confirmed that it is possible to detect the face area with "upward facing state", "sideways facing state", "left / right tilted state" which was impossible with the conventional recognition system, furthermore it showed that nostrils and mouth could be detected.

Chapter 4 describes the application of the detection system of facial feature points in the three-dimensional space constructed in the previous chapter to the control of the meal support manipulator. A spoon was provided at the tip of the three-link manipulator, and the system of the stereo camera developed in the previous chapter was put behind the three-link manipulator. The mouth coordinates of the three-dimensional space detected by the stereo camera was taken as the target position. The opening / closing state of the mouth was detected and used as a trigger to move the spoon attached to the manipulator close to the mouth. When the mouth opened state was recognized, the spoon approached to the mouth of the user, then temporarily stopped in front of the mouth. If the mouth kept opening state, the spoon entered the oral cavity. After the system confirmed that the mouth was closed, the spoon returned to its original position and a series of movements could be done.

Finally, Chapter 5 summarizes the contents from Chapters 1 to 4.

# 学 位 論 文 の 要 旨

論文題目 食事支援装置のための 3次元空間における顔特徴点の検出  
(英訳) Detection of facial feature points in three-dimensional space  
for meal support equipment

氏 名 彭 博 印

本学位論文では、障害者ならびに高齢者の支援を目的として、食事支援装置への適用を前提とした 3次元空間での顔特徴点の検出方法について提案を行っている。

第 1 章では、研究の背景および既存の研究、本研究の位置付けについて記述している。現在、日本では少子高齢化が進み、介護者の不足が懸念されており、労働者人口の不足の問題を解決することがひとつの重要な課題となっている。特に、被介護者の楽しみのひとつである食事の介助は、介護者に依存するところが大きい。介護施設などでは、少数の介護者が多数の食事の介助を行っている場合もある。このような社会的な背景から、これまでも食事支援マニピュレータの開発が行われてきた。しかしながら、食事支援マニピュレータの主な役割は、食べ物を口元まで運ぶことであり、食べ物を口腔内へと入れるためには、使用者自身が顔を移動させる必要がある。そのため、状況によっては使用者に負担を与える可能性がある。そこで本論文では、3次元空間における顔特徴点の位置の検出方法を提案し、食事支援マニピュレータへの適用を行うことを目的とした。

第 2 章では、開発したシステムの概略とステレオカメラによる三次元位置の検出について記述している。本システムでは、2つのカメラを横並びに配置し、得られた 2つの画像から、視差を利用することで、注目するポイントの三次元の位置計測を行っている。それぞれの画角と光軸、2つのカメラの距離を計測し、さらに画像ひずみを校正した結果、三次元空間内において注目するポイントは、誤差が±1%以内で位置の計測が可能であることを示した。

第 3 章では、顔、鼻孔および口の自動検出手法、検証実験の結果について記述している。取得した画像から人物の顔を抽出するため、RGB色空間での肌の色情報について、異なる明るさの環境下での特徴について調べた。その結果、明るさに関係なく、Rの成分は皮膚、鼻孔、唇の部分が他の領域よりも強く、またRとBの成分の差は非常に大きくなる特徴が見られた。この特徴について、RGBそれぞれの値、差、割合の関係に着目して定式化を行い、肌色に適合する条件を見出した。一方、画像処理により顔を検出する場合、顔全体を1つの連結成分として認識することが重要であるが、影や眼鏡などの顔の状態に応じて連結成分として認識されないことが多い。そこで、原画像の大まかな情報を利用する

こととし、画像にモザイク処理を施した。640×480 ピクセルの取得画像に対して、16×16 ピクセルを1ブロックとして、40×30 ブロックに分割し、代表点の平均色から各ブロックの代表色を求めた。その代表色が肌の色の条件を満たしているブロックを抽出し、さらに平滑化処理することで、大まかな顔領域を推定した。推定した顔領域内において、オリジナル画像に対して二値化、ラベリング処理を行い、各連結成分のアスペクト比、面積から鼻孔の候補の絞り込み、鼻孔の検出を行った。また目や口、顎の周辺は暗い領域が多いのに対して、鼻孔周辺は明るい肌の割合が多いことに着目し、顔領域内において肌色の占める割合を求めることで鼻孔周辺領域を推定し、検出された鼻孔位置の確認に利用した。これにより、鼻孔位置の未検出、誤検出を防いだ。得られた鼻孔位置から口周辺の領域を推定し、二値化処理により、口の開閉状態が認識可能なシステムを構築した。被験者による検証実験を行った結果、従来の認識システムでは不可能であった“上を向いた状態”、“横を向いた状態”、“左右に傾けた状態”での顔領域が検出可能であることが確認でき、さらに鼻孔、口が検出可能であることを示した。

第4章では、前章において構築した三次元空間での顔特徴点の検出システムを食事支援マニピュレータの制御に適用したことについて記述している。3リンクマニピュレータの先端にスプーンを設け、前章で開発したステレオカメラのシステムを3リンクマニピュレータの背後に設置した。ここでは、ステレオカメラにより検出された三次元空間の口の座標を目標位置とし、またスプーンを口元に接近させるためのトリガとして、口の開閉状態を利用した。口の開いた状態が認識された場合、スプーンが使用者の口元に接近し、その後、口の前で一旦停止し、さらに口を開き続けることでスプーンが口腔内に入る仕様とした。口が閉じられたことが確認された後は、スプーンが元の位置に戻り、一連の動きが可能であることが示された。

最後に第5章では、第1章から第4章までで得られた知見をまとめ、本論文の総括とした。