

カラー動画画像から手指の動作の認識

7L-4

陳 謙

辻 三郎

イメージ情報科学研究所

大阪大学

はじめに

本研究はヒューマンインタフェースのための手指の動作の認識に関するものである。手の動作の認識の研究は近年、仮想現実感などの研究とともに盛んに行なわれ、多種の認識手法が提案された。しかし、それらの手法の大部分は、認識の時に、手の形状は変化してはいけないとか、LEDなどのものを指先に付けたり、特殊な手袋を使用することのような、ユーザに不自然な制限を加えるものが多い。また、指先の位置だけを求め、その情報だけを利用して手の動作の認識を行なう手法も提案された。

手の動きを不自然に制限したり、手の上に特殊な装置をつけることは、手の動作を認識する装置を、ヒューマンインタフェースとして利用する際、ユーザの創造性を損なう可能性があり、望ましいことではない。また、人間の手の動作は非常に繊細で豊なので、手の指先の位置の情報だけで表現することは不十分である。以上の理由で、我々は手指を重視して、視覚による非接触的な手法で手の動作認識の研究を行なった。

我々は、一般的なオフィスのような自然照明の環境を想定し、カラーテレビカメラで撮像された入力画像を次の処理によって手指の認識を行なう。まず、肌色の情報を用いて入力したカラー動画画像から手と背景を分離し、手の領域を抽出した後、手の領域のスケルトン(skeleton)を求める。人間の手指は太さがほぼ一定の円筒状のものという制約条件を利用して、手の領域のスケルトンから手指の部分抽出し、手指のスケルトンから手指の位置、方向と長さを決定して、手指の動きを求める。

1 カラー画像から手の領域の抽出

手の動作認識の最初の処理は入力画像から手の部分を取り出すことである。この時、色彩情報を利用して、入力画像の手部分と背景の部分とを分離し、手の領域を抽

出する。人間の手の皮膚の色、明るさはほぼ一定で、背景の色や明るさと違うので、肌色の部分を抽出することにより手の領域を求めることができる。

手の領域を抽出する処理をより安定に動作させるため、我々は入力したカラー画像の赤緑青で表現された色彩情報を人間の視覚特性と一致した色空間に変換する。

まず、入力した赤緑青(RGB)画像を次の式を用いてCIEのXYZ色空間に変換する。

$$\begin{cases} X = 0.619R + 0.177G + 0.204B \\ Y = 0.299R + 0.586G + 0.115B \\ Z = 0.000R + 0.056G + 0.944B \end{cases} \quad (1)$$

そして、Farnsworth^[1]が提案した均等色度図表を利用して、入力画像を人間の色彩感覚特性と一致した色空間に変換する。

手の領域を抽出するために、まず肌色の基準点との距離がある閾値以下の画像中の画素を集めて、領域を作る。そして、画像雑音などの影響でできた小領域を取り除くことにより、手の領域を抽出する。

2 手指の抽出

手の動作を認識するために、手のもっとも表現力の高い部分である手指を抽出する必要がある。そのために抽出した手の領域に対して、距離変換を行い、距離変換画像の尾根を追跡することによって、その領域のスケルトン(skeleton)を求める。

このスケルトン上の点の位置は領域の“中心軸”の位置で、その値は領域のその場所での太さを表している。手指の太さはほぼ一定であるという制約条件を利用して、値が手指の太さの範囲以内のスケルトンの点を抽出し、互いに隣接する点を集めて、手指の中心軸になりうるスケルトンの曲線を作る。伸びている手指の長さが一定の範囲内である制約条件を利用して、あまり短い曲線(及び孤立点)と長過ぎる曲線を取り除くことにより手指のスケルトンを抽出することができる。

⁰Hand Motion Recognition from Color Image

¹Qian Chen, Lab. Image Information Sci. & Tech.

²Saburo Tsuji, Osaka University

3 手指の動作の認識

求められた手指のスケルトンを利用して、手指の重心の位置、方向、長さおよび平均の太さを決定する。親指は太くて短く、小指が一番細いので、求められた指の形状情報を利用して抽出した手指がある程度分類できる(親指、小指とその以外の指)。抽出した手指の情報を用いて、手の領域から指の部分を消して、手の平の部分を得られる。この手の平の部分の情報と指の情報を利用して手指がさらに細かく分類できる。

入力した動画の各フレームから手指と手の平の部分を抽出し、フレーム間における手指の対応づけを行ない、各指の重心、長さ及び方向を含む指の形状の変化から、指の並進運動、回転運動の情報を抽出することができる。

各指の運動情報と手の平の部分の重心の位置の変化の情報を利用して、手の全体の運動、姿勢の変化は分析できて、手の運動は理解できる。

4 実験

我々はソニーのHi8の8ミリビデオカメラを用いて、手の動きのシーンを撮影して録画した。普通の机の上に黒い紙を敷いて、その約80センチ上の所にビデオカメラを設置し、手を動かしながら入力画像を撮影した。撮像された画像を256×240の24ビットのRGB画像に変換し、計算機に入れ、認識処理を行なった。我々はこの認識プログラムを表情合成システム及び音声合成機器に連結させ、認識した結果をそれらの機器に送り込み、コンピュータの中の合成人間の喋りで認識結果を出力する試みを行なった。実験の結果、毎秒3-4フレームの速さで処理ができた。

5 おわりに

本論文では我々が提案したカラー画像から手指の動作の認識の方法を紹介した。手の認識において、各指の動きを非接触的に抽出することは非常に重要である。我々は、処理を簡単で且つ確実に遂行するために、色情報を用いて手の抽出を行ない、距離変化でスケルトンを抽出することによって、手指の認識を行なった。実験では、シーンの中の手が確実に抽出でき、手指の動作が確実に求められて、我々の手法が有効であることを実証した。我々はこれから2台のカメラで撮像を行ない、ステレオ画像を用いて手指の完全な3次元運動の抽出を行ない、そしてパイプラインプロセッサを利用することによってリアルタイムの手の動き認識を行なう計画である。

参考文献

1. Günter Wyszecki and W. S. Stiles, COLOR SCIENCE, John Wiley & Sons, Inc.
2. Farnsworth, D. The Farnsworth rectilinear uniform chromaticity scale diagram No. 38. Memorandum Rep. 44-1, New London, Conn. Med. Res. Lab. U.S. Submarine Base, April 1944.
3. Qian Chen and Saburo Tsuji "A Hierarchical Method that Solves the Shape and Motion from an Image Sequence Problem", Proc. IROS'92, pp 2131-2138.
4. Qian Chen and Saburo Tsuji "A Generalized Shape and Motion From Image Sequences", Proc. First Korea-Japan Joint Conference on Computer Vision, pp 343-349.
5. 亀田能成, 美濃導彦, 池田克夫, "シルエットを利用した指の三次元形状推定法", MIRU'92, pp II-239-II-246.
6. K. Ishibuchi, H. Takemura and F. Kishino: "Real-Time Hand Shape Recognition using Pipe-Line Image Processor", Proc. Human Interface Symposium, pp 275-280, 1991.