

5C-1 2段階のモデルを用いたビデオ映像からの人物の抽出とその利用

柳沼良知 影山誠 坂内正夫
東京大学生産技術研究所

1 まえがき

ハードウェアの高速化、記録媒体の大容量化にともない、カラー画像、カラー動画の認識、加工、利用に対するニーズが高まりつつある。このような試みの一つとして、現在、TVのドラマ映像を対象として、一般に多くの動画画像に対して存在するシナリオや音声といった関連情報の認識を協調させながら、ドラマシーンのより高度な認識を行なうシステムの開発を行なっている⁽¹⁾⁽²⁾。このようなドラマシーンの認識、理解、加工、編集のためには、ドラマシーンにおいて重要な意味をもつと思われる「人」の抽出が必要不可欠である。本稿では、これを実現するため、予め、「肌色と黒」という単純なモデルによって人間の大きな抽出を行ない、次に、動き情報をも考慮に入れた2段階のモデルによってドラマシーン中の人物の抽出を行なう方法について述べる。また、抽出した人物情報の利用例として、人物の特徴抽出の実験結果についてもあわせて述べる。

2 2段階のモデルによる人物抽出方法

2.1 「肌色と黒」による人物抽出方法

第1段階の人間モデルとして、画像中から肌色領域(顔)と黒色領域(髪の毛)を抽出することにより、大きな人物の抽出を行なう。

a) 画像のRGBをHVCに変換する。(ここでは、HVCへの変換の際、MTM⁽³⁾と呼ばれるアルゴリズムを用いた。)

b) 色相(H)を $0^\circ - 360^\circ$ とし、 $270^\circ - 350^\circ$ の領域を肌色の領域として抽出する。

c) 明るさ(V)を10段階(0-9)とし、明るさ4以下の領域を黒(髪の毛)の領域とする。

d) 肌色領域の外接長方形の上辺と黒色領域の外接長方形が重なれば、その領域を人間とする。

図1に実際のTV画像からの「肌色と黒」のモデルによる人間の抽出結果を示す。

2.2 動き情報をも考慮した人間モデルによる人物の抽出

より精密な人間モデルとして、動き情報をも考慮した人間モデルを用いて人間の抽出を行なう。

a) 前述した「肌色と黒」という単純な人間モデルにより大きな人間の抽出を行なう。そして、この結果から、人間の体がどこに存在するかを推定する。また、肌色領域の重心位置を比較することで、大きな人間の動きを推定する。実際の画像から人間の体の位置を推定した結果を図2に示す。

b) 人間の胴体と考えられる部分をメッシュに区切り、その近傍で動かし、もう一枚の画像とのパターンマッチングを行なう。そして、それぞれのメッシュがどのように動いたかを求める(図3)。

c) 「肌色と黒」のモデルから得られた大きな人間の動きから、実際の人間の動きの範囲を推定する。図3において、 \times が肌色の外接長方形の重心の動きから推定された人間の動きを示している。その近傍で、最も動きの頻度の高い部分を選びだし、人間の実際の動きとする。図3の場合、x方向に8画素、y方向に3画素動いていることが分かる。

d) 上述(c)で求められた人間の動きを中心として、ある範囲を定め、その動きの範囲に含まれるメッシュを人間の領域として抽出する。

2段階の人間モデルによる人物抽出の結果を図4に示す。

3 抽出した人物情報の利用

ドラマシーン中の、ある特定の人物(例えば、主人公)を、すべて他の人物に置き換えるような場合、映像中の同一人物をすべて選び出すという処理が必要になる。この場合、その人物を他と区別できる「特徴」の抽出が必要不可欠である。ここでは、人物の特徴として、その人が着ている洋服の色の抽出法について述べる。

a) 前述の方法により、人物の抽出を行なう。

b) 抽出された人物の洋服が存在すると思われる部分を推定する。(図5)。

c) 洋服部分について色のヒストグラムを求め、これを服の特徴として抽出する(図6)。

4 むすび

本稿では、ドラマシーンの認識、理解、加工、編集の際に必要な不可欠となる、ドラマシーンからの人物の抽出法について述べた。また、抽出した人物情報の利用例として、人物の特徴抽出の実験結果についてもあわせて述べた。

今後の課題としては、抽出された人物をどのように高度利用していくか、抽出された人物を使って、どのようにドラマシーンの認識を高めていくか、について更に検討を行なう予定である。

文献

- [1] 柳沼良知、影山誠、坂内正夫：協調型ドラマシーン理解システムによるシーン・カット・音声の対応付け(理論)、第45回情報処理学会全国大会、4J-3(1992)。
- [2] 影山誠、柳沼良知、坂内正夫：協調型ドラマシーン理解システムによるシーン・カット・音声の対応付け(実験)、第45回情報処理学会全国大会、4J-4(1992)。
- [3] 宮原誠、吉田育弘：色データ(R,G,B) ↔ (H,V,C) 数学的変換方法、テレビジョン学会誌、Vol.43, No.10, pp.1129-1136(1989)。



Fig.1 「肌色と黒」のモデルによる人間の抽出

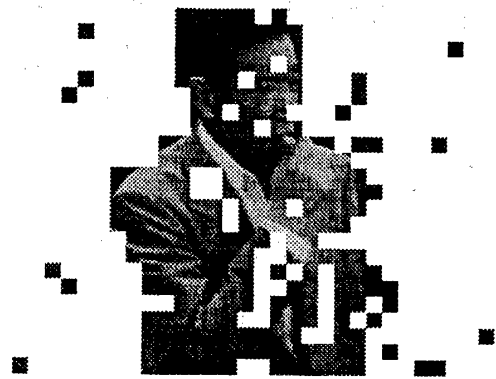


Fig.4 2段階のモデルによる人間の抽出結果

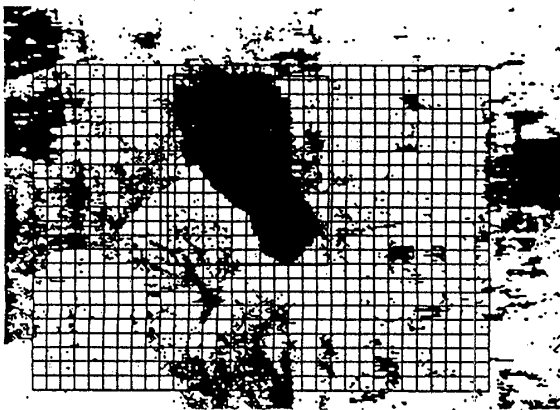


Fig.2 人間が存在する部分の推定

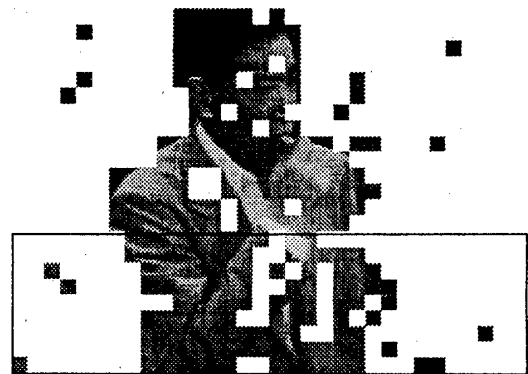


Fig.5 洋服が存在する部分の推定

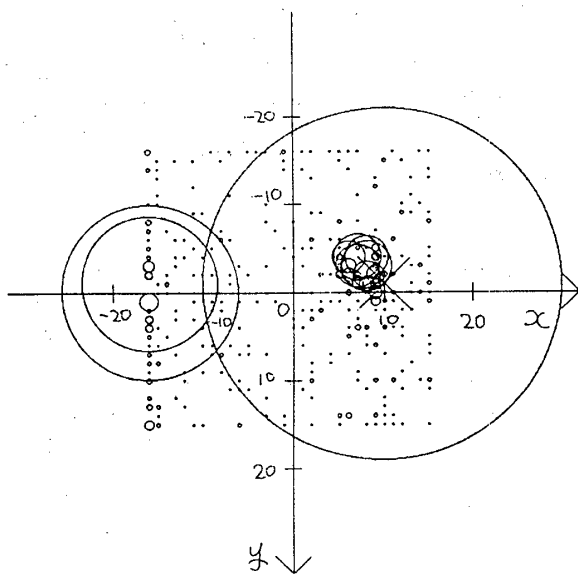


Fig.3 人間領域の各部の動きの推定

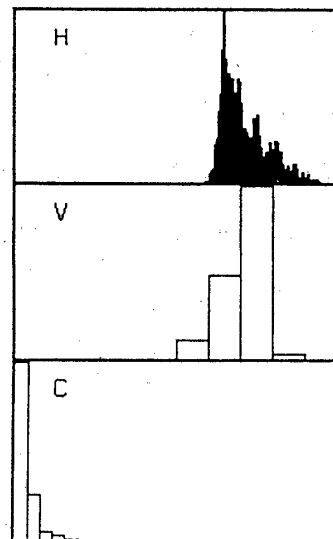


Fig.6 洋服の特徴の抽出結果