

ドラマシーンからの人物の抽出

Human Extraction from the Drama Scene

柳 沼 良 知*・坂 内 正 夫*
Yoshitomo YAGINUMA and Masao SAKAUCHI

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, カラー動画像の認識, 加工, 利用に対するニーズが高まりつつある。カラー動画像のオーサリング¹⁾, カラー画像の検索のためのキーワードの自動抽出²⁾, カラー動画像のシーンの自動分類³⁾等の研究は, それらのニーズに答えるものである。また, ドラマシーンの認識を高度に行うことができれば, 主人公や特定物の抽出, その結果を用いた主人公の置き替え等の加工・編集, ドラマシーン検索のためのキーワード抽出の自動化などが可能になる。

これらを実現する試みの一つとして, 現在, 映像・シナリオ・音声といった対応が必ずしも明確でないメディア情報を統合し, ドラマシーンのより高度な認識を行うシステムの開発を行っている⁴⁾。本システムの特徴としては, メディアの種類や数に依存しない協調システムを実現するため, メディアに依存しない共通の概念を用いて協調を行う点が挙げられる。各メディアには, 概念を抽出するためのアナライザが取り付けられており, 時間・人数・場所等の概念が抽出される。抽出された概念はスーパーバイザによって協調が行われ, また, スーパーバイザによってその協調結果が保存される。スーパーバイザには, あらかじめ, どの概念をどのように協調させれば目的が達成できるかが規定されており, スーパーバイザとアナライザの間は, 共通概念によって会話が行われる。また, 「肌色が一定の明るさになるように照明が調整される」等のドラマシーンの構造を認識に最大限利用しているのも本システムの特徴の1つになっている。現在, 本システムを用いた主人公の抽出についての検討を行っているが, ここでは, 本システムによる主人公抽出の際に必要な, 映像からの人物の抽出, および, 人物の特徴の抽出について述べる。

2. 映像からの人物の抽出

ドラマシーンの加工, 編集, キーワード抽出のためには, 人物の抽出, 特に, そのドラマにおいて頻繁に登場する, 主人公の抽出が重要である。ここでは, 映像中から肌色領域と黒色(髪の毛)領域を抽出することにより, 人物を抽出する方法について述べる。本システムは, Sparc Station 上に, C 言語を用いて作成された。実験に使用したドラマシーンは, 一般に放映されているものを使用し, 映像の約7分20秒分のうち, シーンチェンジとして抽出された22枚の画像に対して処理を行った。人物の抽出は, 以下のように行った。

a) 人間の顔の部分を抽出するために, 映像から肌色領域の抽出を行う。このため, 画像のRGBをMTM⁵⁾と呼ばれるアルゴリズムによりHVCに変換した。色相(H)は, 0°-360°とし, 270°-350°の領域を肌色領域の候補として抽出した(図1)。これは, 大まかには, 赤色の領域に対応する。ノイズの影響を避けるため, 8近傍で連続する領域の面積が400画素以上の領域を肌色領域とした。図1では, 肌色領域, および, それを囲む外接長方形が示されている。ドラマシーンでは, 照明を登場人物の顔色が自然に映るように調整するため, 肌色の領域については, このパラメータにより, 良好に抽出する

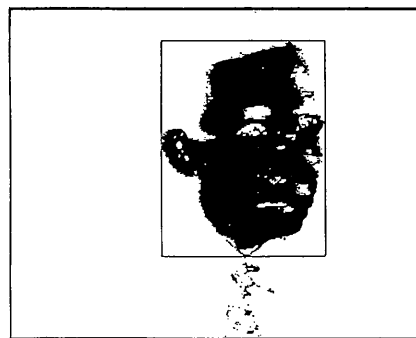


図1 肌色領域の抽出

*東京大学生産技術研究所 第3部

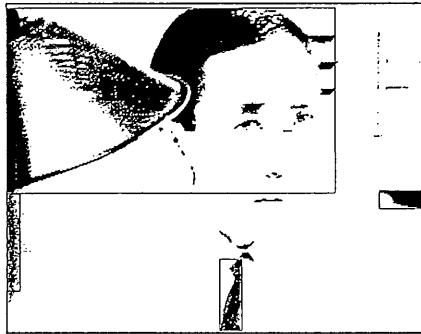


図2 黒色(髪の毛)領域の抽出

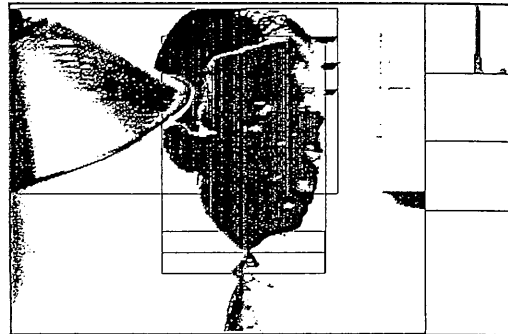


図4 服の色の抽出

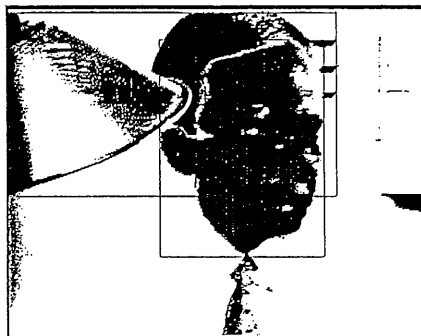


図3 人物の抽出

ことができた。

b) 黒色(髪の毛)領域を抽出するため、明るさ(V)を10段階(0-9)とし、明るさ4以下の領域を黒色領域の候補とした(図2)。ノイズの影響を避けるため、8近傍で連続する領域の面積が400画素以上の領域を黒色領域とした。図2に抽出結果を示す。図2では、4つの黒色領域、および、それを囲む外接長方形が示されている。黒色領域は、影などの影響により、肌色領域に比べて、ノイズが多いことが分かる。また、実験に用いた画像では、髪の毛と、その手前にある黒い照明器具が重なっているため、これらの領域が1つの黒色領域として認識されている。

c) 肌色領域の上に、黒色領域が存在すれば、それを人であるとする。具体的には、肌色の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、その領域を人間とした(図3)。

22枚の画像について実験を行った結果、この方法で、約6割の人物の抽出が行えた。

3. 人物の特徴の抽出

ドラマシーン理解システムによる主人公の抽出の際、

映像中に複数の人物が存在する場合、それぞれの人物の区別を行う必要がある。ここでは、人物の区別を行うための人物の特徴として、服の色の抽出を行った。具体的には、肌色領域の外接長方形の下部分(面積1/5)を洋服部分とした(図4)。この部分の色相(H)のヒストグラムを洋服の色とする(図4右上のヒストグラム)。もし、ある画面上に存在する複数の人物が、それぞれ異なる色の洋服を着ているならば、この服の色のヒストグラムを比較することで、人物の区別を行うことができる。服の色が一致するかの判断は、ヒストグラムにより、

$$\chi^2 = \sum_{i=1}^n \frac{(H_2(i) - H_1(i))^2}{H_1(i)} \quad (1)$$

を求めることを行うことができる。 $H_1(i)$ 、 $H_2(i)$ は、それぞれ、比較する服の色のヒストグラムを表す。もし、 χ^2 が十分小さければ、同一の服の色と判断する。

4. む す び

映像、シナリオ、音声等の情報を統合し、ドラマシーンのより高度な認識を行うドラマシーン理解システムにおいて、主人公抽出の際に必要な、映像からの人物の抽出、および、人物の特徴の抽出について述べた。ここでは、映像中から肌色領域と黒色(髪の毛)領域を抽出することにより、人物の抽出を行った。また、人物の区別のために必要な、人物の特徴として、服の色の抽出もあわせて行った。今後、共通概念を用いたドラマシーン理解システムによって、具体的に何が出来るか、また、多数のメディアの協調をどのように実現するか等について検討を行っていく予定である。(1992年9月1日受理)

参 考 文 献

- 1) 上田博唯, 宮武孝文, 吉澤 聡: "認識技術を応用した対話型映像編集方式の提案", 信学論, J75-D-II, 2, pp. 216-225 (1992)
- 2) M. Sakauchi and J. Yamane: "Realization of fully automated keyword extraction in image database system",

研 究 速 報

SPIE International Symposium on Optical Applied Science and Engineering (1992)

3) Y. Gong and M. Sakauchi: "A Method for Color Moving Image Classification Using the Color and Motion Features of Moving Images", ICARCV'92 (1992)

4) 影山 誠, 柳沼良知, 坂内正夫: "マルチメディア協調型ドラマシーン理解システムによるカットとシーンの対応付け", 1992年電子情報通信学会秋期大会, (1992)

5) 宮原 誠, 吉田育弘: "色データ (R, G, B) ↔ (H, V, C) 数学的変換方法", テレビ誌, 43, 10, pp. 1129-1136 (1989)