

研究課題別研究評価

1. 研究課題名: 人物行動を伝えるための知的映像撮影と編集

2. 研究者名: 中村 裕一

3. 研究のねらい:

映像を誰でも手軽に使えるコミュニケーション手段とするためには、映像撮影の問題を見直し、それをサポートするシステムを用意することが必要である。また、TV や映画のような形態にとらわれずに、受け手にわかりやすく、短時間で利用できる形で提供する方法を探ることが必要である。

この研究ではその一つのアプローチとして、料理や組み立て等の解説(プレゼンテーション)場面を題材とし、カメラマンの機能(人間の行動を知的に撮影する)、ディレクターの機能(人間の行動を認識して映像を知的に編集する)、マルチメディア・デザイナーの機能(映像が含んでいる情報をわかりやすく提示する)の3つの観点から、映像によるコミュニケーションをサポートする方法論を考え、それを検証するためのプロトタイプシステムを構築することを目的とした。

4. 研究結果及び自己評価:

研究結果

上にあげた3つの機能を実現するために、各々次のような手法を提案し、その検証実験を行ってきた。

カメラマンの機能(人間の行動を知的に撮影する): 顔や手先など、撮影の主対象となる部分を複数のカメラで常に追跡して、いつでも映像として利用できる状態にする自動化撮影システムを構築した。何をどのように伝えるかという目的とカメラの自動制御アルゴリズムやパラメータとの関係を探り、不快感がなくわかりやすい映像を取得する方法を提案した。

ディレクターの機能(人間の行動を認識して映像を知的に編集する): 人間の行動(ここではプレゼンテーションを対象)において重要な意味を持ち、注目する必要がある場面や部分を検出する手法を提案した。注目すべき部分は、時間的・空間的に常に変化するため、人間の行動(体の動きや発話等)を利用して、もっとも見せたい部分を検出するが重要なポイントである。また、シナリオ等のあらかじめ用意された情報を用いることも有効であり、シナリオの記述と人間の言動を照合する手法も提案した。

マルチメディア・デザイナーの機能(映像が含んでいる情報をわかりやすく提示する): プレゼンテーションのような目的がある場合、各時点の人間の行動やその他の被写体には意味や因果関係がある。それらをうまく記述して、映像をわかりやすく提示するような映像提示手法を提案した。映像中のフレーム(画像)をアイコンとし、それを概念図のように並べることで映像の構造を表現する、ビデオアイコンダイアグラムの手法を提案した。

このような考え方に基づいて、映像の自動撮影・編集システムを構築した。このシステムでは、位置センサや画像処理により話し手や特定物などの位置を取得し、複数台の首振りカメラを制御すること

で自動撮影を行う。各々のカメラで撮影された映像は、MPEG エンコーダを通して保存され、ランダムアクセスが可能になる。また同時に、位置センサと音声認識を併用して話し手の動作認識を行い、映像へのタグ付けを行う。さらに得られたタグを基にして、視聴者が見たいと思う部分を効果的に提示する。実際に撮影実験を行った結果、カメラの切り替えを含め、自然な映像が得られることを実証した。このシステムには、実用的なシステムとしての種々の応用が考えられると同時に、映像撮影や映像編集を実証的に探るためのツールとしての利用が考えられる。これから、更に発展させていく予定である。

自己評価

「さきがけ 21」プログラムの支援を受け、3 年間最も力を注いだのは、システムを実際に動作させることであった。原理的には簡単な処理の組合わせではあるが、実際に総合システムとして構築し、実時間で動作させたのは世界で初めてである。心配された映像の「質」も、時には素人ビデオ以上の結果が得られるところまでいったことに満足している。一般的な問題を扱うためには、まだまだ不十分な点が多く、すぐに実用化ができるものではないが、映像コンテンツ取得を自動化するシステムの可能性を示した点で、十分な意義があった。

このように実証的なシステムを構築したことから、他の研究との協力が行えるようになった。科学研究費・特定領域研究「視覚情報メディアのためのパターン認識・理解」では、複合コミュニティ空間(現実世界と仮想世界を合成した場で複数の人がコミュニケーションを行う空間を呼ぶ)をサポートするシステムとして利用されている。また、本年度から 5 ヶ年計画の科学研究費・学術創成研究「人間同士の自然なコミュニケーションを支援する知能メディア技術」において、様々な用途に用いることのできる映像コンテンツを獲得するためのプラットフォームとして活用される予定である。

今後は、扱う対象を少しずつ広げること、人間の自然なさりげないコミュニケーションを扱えるようにすること、マルチメディアコンテンツとしての利便性を高めること等、種々の点から高度なシステムへ向けてのアプローチをする予定である。さらに、このような実際的な研究とともに、映像を情報としてどのように表現し、扱うかという点について考察していきたい。映像がコミュニケーション手段として日常に深く浸透しているにもかかわらず、その情報学的な表現や扱いにはまだ十分なアプローチがなされていないのが現実である。映像がこれからますます大量に流通・蓄積されていくことを考えると、映像の本質を「情報と知」の面から探ることが必須であると考えている。

5. 領域総括の見解:

映像や映像を基にしたマルチメディアコンテンツがコミュニケーションの手段として重要性を増す中で、誰でもが手軽に使えるコミュニケーション手段にすべく、総合システムを開発し、世界で初めて実時間で動作させた成果は高く評価できる。また、応用範囲が広い技術であり、今後の研究発展に期待できる。

6. 主な論文等:

* (1) Y.Nakamura, M.Ozeki, Y.Ohta, An Intelligent System for Capturing Presentation on Desktop Manipulations --- Supporting for Video Contents Production, Chapter2 in Internet Based Teaching and Learning, R.Howlett and L.Jain eds., Advanced Information Processing Series,

Springer-Verlag (accepted)

(2) M.Ozeki, Y.Nakamura, Y.Ohta, Camerawork for Intelligent Video Production --- Capturing Desktop Manipulations, Proc. Int. Conf. on Multimedia and Expo, TA1.5, pp.41--44, 2001

(3) M.Murayama, Y.Nakamura, Y. Ohta, Diagram Generation From Tagged Texts Toward Document Navigation, Proc. Int. Conf. on Multimedia and Expo, FP0.4, pp.716--719, 2001

(4) * 村山正司, 伊津野英克, 中村裕一, 大田友一, ビデオアイコンダイアグラムによる映像内容の構造表現, 信学技報 PRMU2001-45, pp.47--54, 2001

(5) Y. Nakamura, Multimodal Approach toward Intelligent Video Production, International Workshop on Multimedia Intelligent Storage and Retrieval Management, pp.1-8, 1999