

## 多視点ハイビジョン映像生成システムの開発

映像情報メディア学会誌, Vol. 64, No.4, pp.622-628 (2010)

富山仁博, 岩館祐一

近年、映画やドラマ、CMなどにおいて、被写体を取り囲むように配置した複数のカメラで被写体を撮影し、その映像をカメラの並びに沿って切り替えて表示する多視点映像表現が盛んに用いられている。この多視点映像表現をスポーツ中継で利用するための多視点ハイビジョン映像生成システムを開発した。スポーツ中継で使用する多視点映像生成システムはセッティングが1日程度で完了し、競技中の選手の動きや競技の進行に合わせて流動的に使用できなければならない。そこで、位置を固定した12台のハイビジョンカメラと3台のPCによる小規模な構成とし、位置を固定したカメラで撮影した映像から選手が映像の中心になるように仮想的にパン・チルト・ロール回転ができる射影変換手法を考案した。この射影変換処理をGPUを用いて高速に実行し、多視点映像の編集・送出手動化したので、競技終了直後から1分以内に多視点映像を送出することができた。開発したシステムは、2006年からNHKの体操競技の中継等で解説映像として利用された。

## 2次元コードを利用した情報連結方式

情報処理学会論文誌, Vol.51, No.11, pp.2081-2088 (2010)

川喜田裕之, 西本友成, 井上友幸

AV機器などの機器情報（例えば、機器IDなど）を2次元コード（QRコード）としてディスプレイ等に表示し、複数の機器で表示されたQRコードを連続して読み取り、QRコードリーダー上で各情報を連結する方式を開発した。QRコードの仕様では、複数のQRコードを続けて読み取り、符号化された各情報を連結する機能（構造的接続）が規定されている。この機能は、あらかじめ複数のQRコードに分割された固定長の情報を読み取った後、連結して元の情報に戻す機能であり、任意の情報を連結することはできない。そこで、QRコードの符号化に必要なパリティ値が連結後も変化しないような特定の情報（文字列）を可変長の情報に付加し、任意の情報の連結を可能とする方式を開発した。本方式を用いたQRコード生成プログラムを複数のAV機器に実装し、QRコードリーダーとして市販の携帯電話を使い、任意の情報が連結できることを確認した。本方式は携帯電話を制御端末として利用し、複数のAV機器に制御信号を同時に送信する際の符号化などに適用できる。

## Baseball Digest Production System using Inductive Logic Programming

Journal of Multimedia, Vol.5, No.4, pp.343-351 (2010)

宮崎 勝, 柴田正啓, 八木伸行

放送局では夜のスポーツニュースなどでその日に行われたスポーツの試合のダイジェストを放送することが多い。また、データ放送などで野球の試合のメタデータ配信も一般的に行われており、メタデータを利用して受信機側でダイジェスト映像を生成できる可能性も高くなってきた。そこで、帰納論理プログラミング（論理プログラムの形式で記述された多くの具体的な事例から目標とする概念定義を導き出す研究分野）の技術を利用して、手作業で制作したダイジェスト映像のメタデータからハイライトシーンの概念定義を学習する手法を提案し、決定木学習（事例データから目標を分類する木構造を作る機械学習手法）との比較でその優位性を示した。番組制作者が持つノウハウをオントロジー（概念間の関係を記述したもの）として獲得する枠組みを実現することで、ハイライトシーンの候補となる複雑な試合状況を容易に抽出し、ダイジェスト映像の制作に利用することも期待できる。

## Learning Speech Variability in Discriminative Acoustic Model Adaptation

IEICE Trans. Inf. & Syst., Vol.E93-D, No.9, pp.2370-2378 (2010)

佐藤庄衛, 奥 貴裕, 本間真一, 小林彰夫, 今井 亨

字幕放送の拡充や番組メタデータの効率的な制作を目的として、音声認識精度を改善する方法を提案した。これまでの音声認識では、発話内容が決まっていなかった自由な発話の不明りょうな発声（発声変形）部分で誤認識が起き、対談番組などでは十分な認識精度が得られていなかった。そこで、対談番組に特有の言い回し（単語）で発声変形の頻度が高いことを確認し、それらの単語に専用の音響（音素）モデルを識別的に学習する方法を提案した。提案法では、音響モデルの学習用の対談番組から頻度の高い単語を調べ、それらの単語の対立仮説（誤り仮説）から正解単語の信頼度を求める。この信頼度が低い単語にだけ発声変形用の音響モデルを導入し、それらを他の音素と識別するように学習する。報道系情報番組を対象として認識実験を行い、高頻度単語のすべてに専用の音響モデルを追加して学習する従来法と比較した。提案法では、従来法の約半数の音響モデルの追加で、専用の音響モデルをすべてに追加した従来法とほぼ同じ12%の認識誤りを削減でき、発声変形による認識誤りを効率的に削減できることを示した。