3　ユニバーサル放送サービス

デジタル放送やインターネットなどの情報基盤の拡充・普及によって、視聴者には多岐にわたるコンテンツが大量に届けられるようになった。また、視聴者の側でも、多様化した生活スタイルのさまざまな場面に合った方法で、必要なコンテンツを受け取り、楽しみたいという要求が高まってきた。さらに、社会の高齢化や国際化の進展に伴って、視聴者の視覚・聴覚の身体的な条件や言語などの文化的な条件はより分散したものとなってきている。そこで、情報技術などを使って所望のコンテンツをその人に適した形式で提供するための研究を進めている。

人にやさしい情報提示については、開発した情報バリアフリー受信提示システムの長期間の使用評価実験を視覚障害者を対象として行ったほか、触力覚を使って2次元または3次元の情報を提示するシステムの試作と認知的特性の評価を継続した。また、日本語テキストから手話CGへの自動翻訳技術については、気象情報を対象として、翻訳可能なテキストの構文やCG表現の範囲を広げる研究開発を進めた。

字幕制作のための音声認識では、話題が多様に変化するワイドショー番組を対象とした研究に着手した。また、ニュース番組の字幕制作設備のために開発した技術を実運用システムの設備整備に適用した。

高齢者のための音声・音響信号処理では、受信機側で聞きやすい番組音に調整するための背景音抑圧法や音声強調法について開発を進め、有効性を確認した。

言語バリアフリーサービスのための言語処理については、国内外の外国人に向けたサービスを拡充するために、ニュースのやさしい日本語への変換支援技術の研究とニュースの日英統計翻訳技術の研究を進めた。また、視聴者から寄せられる番組への意見を分析するための手法を高度化するための研究を進め、放送現場での分析システムの試作に貢献した。

コンテンツ検索・推薦技術については、ウェブ上にある大量のテキストから単語間の関係を含む意味ネットワークを獲得して番組の検索などに活用する手法の開発を進めた。また、素材映像の検索・再利用のための画像処理・情報抽出、番組検索、視聴状況に基づく番組推薦の研究開発を、メタデータ制作技術を基盤として進めた。その成果を国際標準化やウェブでの実験サービスとして公開した。

視聴者の心理状態推定技術として、番組が視聴者にどのように見られ、どのような心理的影響を与えるのかを、脳活動および視線を手がかりに解明する研究を行い、スーパーハイビジョンの見られ方やその要因、「笑い」と脳活動との関係を探る検討などを進めた。

3.1　人にやさしい情報提示

視覚や聴覚に障害のある人が放送を楽しむことができるようにするために、人にやさしい情報環境を実現する研究に取り組んでいる。

■デジタル放送の情報バリアフリー技術

視覚障害者用の情報バリアフリー受信提示システム(VIA-TV)については、盲ろう者や視覚障害者を対象に長期間のモニター評価実験を行った。評価結果を基に操作性の課題を抽出し、さらなる改善を図った[1]。また、VIA-TVで表示されるメニューや抽出した図表の一覧を、微小な可動点をマトリックス状に配置した触覚ディスプレイに表示し、インタラクティブに操作できる装置を開発した。

■2次元・3次元情報の触力覚提示技術

2次元情報の触覚提示では、グラフィカルな情報を効率的に把握できるようにするために、触覚ディスプレイ上の個別のオブジェクトを異なる周波数で振動させる方式を開発した（図1）。こ

図1　振動提示が可能な触覚ディスプレイ（VIA-TV対応）

の方式は、視覚障害者が地図上の目標の位置やメニューの分類などを迅速に探索することに役立つ。また、触覚ディスプレイで大きな図表を触知する際の操作性を改善するために、フリックやドラッグによる操作性と従来のキースイッチによる操作性の差異を、探索に要する時間と主観評価で明らかにした。この研究の一部は、東京女子大学、KGS（株）と共同で実施した。

3次元の触覚情報の研究では、複数の指にアクチュエーターで力覚を与えて仮想物体を提示した場合と、実物体に触れた場合の形状認知特性の差異の評価実験を、視覚障害者を対象に進めた。仮想物体表面の不連続な箇所で指先が物体に引きつけられるような力を与える方法と、指先以外に手のひらにも仮想物体からの反力を提示する方法を比較した結果、両者を同時に提示した場合に最も形状認知に効果があること、手のひらの一部だけに反力を与える提示では、形状が歪（ゆが）んで感じられることが明らかになった[2]。

■気象情報の手話CG翻訳技術

聴覚障害者用の手話放送サービスの拡充を目指して、日本語テキストから手話CGへの自動翻訳技術の開発を進めている（図2）。2010年度までに開発した、日本語ー手話対訳辞書システムを応用して日本語の文から対訳用例に基づき手話CGに翻訳するシステム[3]では、既知の用例と全く同じ構文の入力しか処理できなかったが、2011年度はこれを拡張し、気象情報を対象に、用例翻訳と統計翻訳を融合し、入力した日本語文を節・句に分けて翻訳を行う技術を開発した。さらに、取り扱える話題の分野は気象情報に限られるが、任意の文を翻訳できるシステムを試作した。一方、手話では手や指の動作とともに、顔の表情や口型などによる非手指動作が重要である。そこで、手話通訳者の顔のモーションキャプチャーデータの収集を継続するとともに、これに基づいて動作するTVML（TV program Making Language）用の人物モデルの開発を進めた。

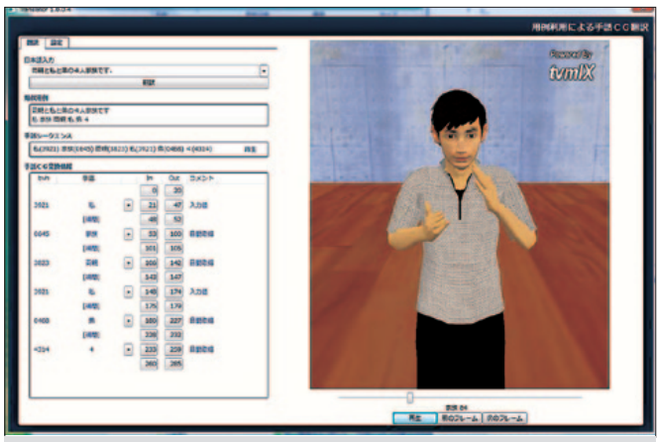

図2　用例翻訳技術による手話CG生成システム

〔参考文献〕
(1) 坂井，半田，大河内，伊福部：“視覚障害者向けデジタル放送受信機とUIの開発評価，” 信学会HCGシンポジウム，HCG2011-B3-4，pp.154-159（2011）
(2) T. Handa, T. Sakai and H. Shinoda: “Evaluation of Size and Shape Perception in Multi-Finger Haptic System,” IEEE World Haptics Conference 2011, pp.523-528（2011）
(3) 加藤，金子，井上，梅田，比留間，長嶋：“用例利用による日本語ー手話CG翻訳システム，” 信学会HCGシンポジウム，HCG2011-I-1，pp.298-303（2011）

3.2　字幕制作のための音声認識

聴覚に障害のある方や高齢の方をはじめ、多くの方が楽しめる字幕放送の拡充を目指して、リアルタイム字幕制作のための音声認識の研究を進めている。

■情報番組の字幕制作

大相撲などのスポーツ番組では、字幕キャスターが言い直した音声を認識して文字にするリスピーク方式の字幕制作システムを、すでに実用化している。2011年度は、総合テレビの情報番組「あさイチ」の字幕放送を開始した。しかし、話題が多様に変化するこのようなワイドショー番組では、話題に追従して言語モデル（単語列の出現確率）を使い分けるなど、今後、いっそうの改善が望まれる。

そこで、料理や園芸といった話題の種類を音声認識結果から随時推定し、あらかじめ適応学習しておいた言語モデルを選択的に使用する音声認識手法を開発した[1]。話題の推定には、従来の統計的手法のほかに、より高い効果が期待できる非負値行列因子分解に基づく手法も検討した[2]。また、字幕キャスターの交代時に、誤り修正結果を利用してより効果的に音響モデル（母音・子音の周波数分布）を適応学習する手法[3]や、認識誤り傾向を効率よく反映する認識手法も開発した。これらの改善研究の結果、「あさイチ」から評価用に選定したリスピーク音声の認識率を88.0%から91.8%へ3.8ポイント改善した。

図1　東日本大震災ニュースでの字幕制作

■ニュース番組の字幕制作

アナウンサーや記者の番組音声を直接認識するダイレクト方式とインタビューなどを対象にしたリスピーク方式を組み合わせたニュース字幕制作システムの開発を進めている。2011年度は、16時ニュースなどの字幕放送でこのシステムを実用化するための設備整備に貢献したほか、東日本大震災では試作機を臨時で放送センターに設置して、ニュース番組の一部で字幕放送を実施した（図1）。

〔参考文献〕
（1）本間，小林，奥，今井亨：“情報番組向け話題スイッチング言語モデルの検討，”音講論（秋），3-P-9（2011）
（2）藤田，奥，小林，今井亨：“非負値行列因子分解に基づく話題性を反映した言語モデルの構築，”音講論（春），3-P-11（2012）
（3）奥，藤田，小林，今井亨：“認識誤り修正部分の重点的な話者適応化手法，”音講論（春），1-P-20（2012）

3.3　高齢者のための音声・音響信号処理

背景音に妨害されて高齢者には聞き取りにくくなっているセリフやナレーションを聞きやすくするための番組音調整技術の研究を行っている。

これまでに、ナレーションと背景音の「耳で感じる音の大きさ（ラウドネス）」の差を逐次算出し、ナレーションの聞きやすさを判定する番組制作支援装置の開発を進めてきた。2011年度は、番組制作現場で運用可能な装置を製作し、スタジオに試験的に導入した（図1）。番組音声担当者と連携し、実運用下での支援効果について検証を開始した。

一方、受信機側で聞きやすい番組音に調整するための背景音抑圧法や音声強調法についても検討を進めた。背景音抑圧については、音声/非音声を自動推定し、その結果に応じて抑圧処理を切り替える手法を考案した。抑圧効果に関する比較実験を行った結果、音質劣化が小さく、90%以上の割合で元の音声よりも背景音が小さくなったと評価され、手法の有効性が確認できた[1]。

また、音声強調については、加齢による聴力劣化の特徴を考慮して、ノイズ環境下での音韻の特徴を、よりはっきりさせるための周波数成分の強調方法を検討した。さらに、音声の内容理解を助ける韻律についても、実時間で処理可能な周波数変化幅拡大手法を考案し、雑音下での高齢者の聞き取りやすさの改善に有効であることを確認した[2]。

図1　高齢者のナレーションの聞きやすさを客観評価する装置

〔参考文献〕
（1）小森，今井篤，清山，田高，今井亨，本間，都木：“音声/非音声区間切替による番組背景音抑圧処理法の検討，”信学技報，Vol.111, No.225, SP2011-66, WIT2011-48, pp.77-82（2011）
（2）田高，清山，小森，都木，今井篤：“音声の聞き取り易さ改善のための韻律調整方法の評価，”日本音響学会聴覚研究会資料，Vol.41, No.9, H-2011-125, pp.699-703（2011）

3.4　言語バリアフリーサービスのための言語処理

■日本語変換支援技術・日英翻訳技術

国内に在住する外国人のためのニュースの「やさしい日本語」変換支援技術と国内外の外国人のための日英翻訳技術の研究を行っている。

やさしい日本語の研究では、通常の気象災害ニュースとこれをやさしい日本語に書き換えたニュースを約40名の中国人日本語学習者に提示して、内容を問うテストを行った。その結果、特に、初級レベルの学習者に対してやさしい日本語の効果が高いことが分かった[1]。

人名、地名、書き換えが難しい専門用語などは、やさしい日本語のニュースにも使わざるを得ない。そこで、これらを用語集や画面表示の工夫でわかりやすくしたウェブのサービスイメージを検討し、技研公開で展示した（図1）。画面右のやさしい日本語のニュースでは、全ての漢字にふりがなが付けられている。また、人名と地名はカラーで表示されている。さらに、下線の付いた語をクリックすると左にやさしい日本語で書かれた説明が表示されるようになっている。

このサービスを実現するためには、日本語教師がニュースをやさしい日本語に書き換える作業と、記者がその内容を確認する作業が必要である。この作業を両者が連携して短時間で行うための「書き換え支援エディター」の検討を進め、その支援機能と

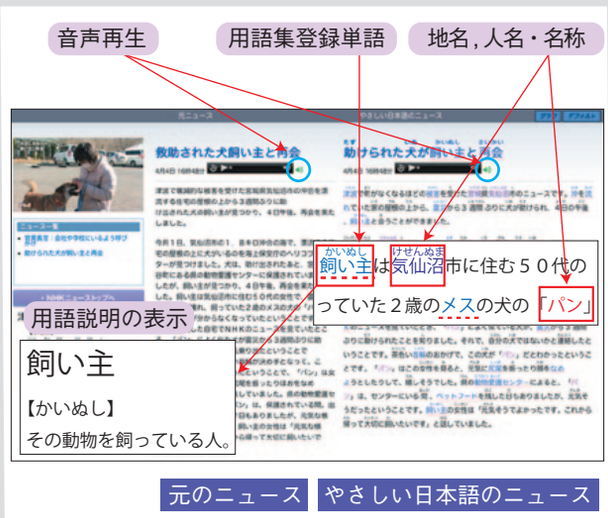

図1　やさしい日本語ニュースのサービスイメージ

して、記事中の難しい単語や、書き換え前後の記事の難易度指標を表示する機能などを開発した。

日英自動翻訳の研究では、ニュース原稿を対象とした統計翻訳技術の研究を進めている。統計翻訳では、日本語と英語のテキストの表現が互いに翻訳として現れる確率を示す「翻訳モデル」の学習が必須である。しかし、日英ニュース原稿は、記事中の全ての表現が相手の言語に翻訳された「直訳」ではなく、文間の対応関係も明確ではないので、既存手法の適用は難しい。これを解決するために、翻訳モデルに加えて、各言語の表現の確率モデルも同時に学習する手法を研究している。2011年度は、句（連続した単語列）を表現の単位として、隣接した句をまとめるための文法を翻訳モデルと各言語の表現の確率モデルに導入し、より長い句の翻訳モデルを学習できるようにした。

■評判分析技術

放送番組に寄せられる視聴者の意見を番組制作に生かすための評判分析技術の研究を進めている。この研究は、意見種別の自動推定のほかに、意見の対象も自動検出することが特徴である。この技術を実現するためには、述語の解析技術を向上させる必要がある。2011年度は、格助詞を手がかりに、述語の対象を推定する技術の研究を進めた[2]。例えば、「いらいらする」という述語の場合、「が」「に」「で」という助詞を持つ単語（例えば、「渋滞で」）が「いらいらする」の対象になりやすい。このような関係を記述した辞書と、述語が受動態で使われた場合などの例外に対応するルールを作って、述語の対象を推定する手法を開発した。

放送現場では簡易投稿サイト（ツイッターなど）への投稿メッセージを収集・分類して業務に生かす試みが始まっており、その自動分類技術が求められている。そこで、実際に業務に使われている分類を参考に「ポジ意見、ネガ意見、中立意見、要望、期待」など、20種に分類する正規表現[注]集を実験で定め、放送センターで開発・運用されている投稿メッセージ収集システムに自動分類機能を追加した。

（注）文字列を照合するために使用するパターンの簡潔な表現方法

〔参考文献〕
(1) 田中, 美野: “「やさしい日本語」ニュースの理解度テスト –ニュースのための「やさしい日本語」の設計に向けて–,” 信学技報, Vol.111, No.228, NLC2011-22, pp.1-6 (2011)
(2) 小早川: “放送番組に対する意見マイニングと述語の解析,” 信学技報, Vol.111, No.119, NLC2011-16, pp.85-88 (2011)

3.5　コンテンツ検索・推薦技術

NHKが保有する映像資源の利用を促進するために、番組検索・推薦のためのテキスト処理技術、画像処理技術、映像再利用のための情報取得・抽出技術、番組検索・推薦技術および視聴状況取得技術の研究を進めた。

■番組検索・推薦のためのテキスト処理技術

2011年度は、2010年度までに開発した番組検索手法を改良し、テキスト構造化を利用した検索・類似性評価技術の研究を進めた。

NHKオンデマンドの番組概要文を類似番組検索に利用する場合、その表現の違いから類似性を正確に求めることができない場合がある。例えば、「地震」と「震災」は類似した単語であるが、単語の表記だけを利用した類似性の計算手法では別のものとして扱われてしまう。このような問題を解決するために、2010年度まで共同研究を進めた（独）情報通信研究機構が開発したツールを活用して、ウェブ上にある大量のテキストから単語間の関係（類似関係、因果関係、上位下位関係、属性関係など）を獲得し、350万語規模の意味ネットワークを構築した。また、意味ネットワークを可視化して手作業で修正できる編集システムを開発した（図1）。この意味ネットワークを利用した類似番組検索システムを試作して意味ネットワークの有用性の評価を進めた。さらに、2010年に開発したWikipediaの変更履歴を利用した類似番組検索技術を改良し、社会で注目されている事象に関連する番組の順位を上げる手法を開発した[1]。

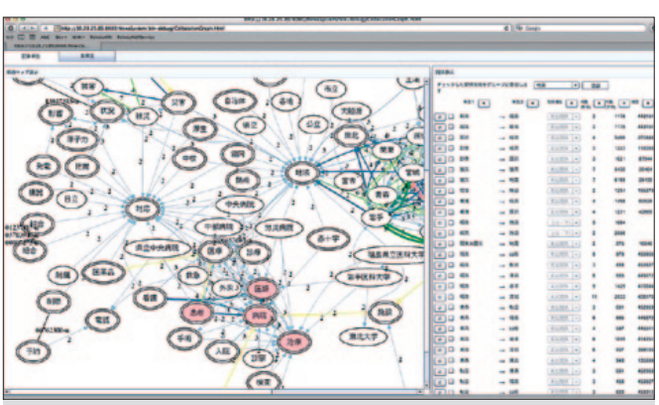

図1　意味ネットワーク編集システム

ニュースの検索精度の向上や効果的な情報提示を目指し、ニュースの構造化の研究を進めている。 構造化は全体の概略的な情報を有しているNHK年鑑と、詳細情報であるニュース原稿やテレビ番組とを対応付けることで行う。2011年度は、EPG（電子番組ガイド）を使うことでテレビ番組を年鑑の記述と対応付ける研究を行った。また、この対応付けを利用し、2010年度に試作したニュース原稿検索システムを番組検索システムに拡張した。

■番組検索・推薦のための画像処理技術

2011年度は、映像中のオブジェクトやイベントの認識、映像の分類および文字の検出に関する研究を進めた。

画像フレーム内の局所領域間の位置関係や、画像特徴量の共起を考慮することで、従来よりも高精度に被写体を認識できる技術を開発した[2]。また、映像中に出現する代表的な画像片（ブロック領域画像）のヒストグラムに注目し、素材映像中のシーンを内容に基づいて分類する技術を開発した。これらの技術を組み合わせて、震災関連などの膨大な素材映像へメタデータを付与する作業を支援するシステムを試作した。さらに、映像中の文字情報を活用するために、L字画面[注]に提示されるテキストを検出・認識する技術を開発した。番組映像の冒頭から順次L字画面の出現と提示テキストの変化を検出し、OCR（光学式文字読取装置）処理することで、L字画面に提示されたテキストをメタデータとして蓄積することが可能となった。

■映像再利用のための情報取得・抽出技術

素材映像の再利用を促進するために、映像の加工・合成が容易にできる情報の取得・抽出技術の研究を進めている。2011年度は、撮影映像からカメラ姿勢情報（位置や向き）を推定する技術の改善と被写体の抽出などの処理を容易にするための領域分割手法の研究を進めた。カメラ姿勢情報の推定技術に関しては、従来手法と比較して11％の精度向上を実現するとともに、画像特徴の抽出条件を動的に調整するアルゴリズムの開発などを行い、推定性能の頑健性を向上させた。動画像の領域分割では、3次元時空間データを扱う必要があるが、2次元画像で一般的に用いられている領域分割手法のMean-Shift法をそのまま適用すると膨大な処理時間がかかる。そこで、3次元時空間データを並列分散処理に適するように手法を改良し、処理時間の高速化を図った。また、簡易な指示で所望の被写体を抽出するためには領域を被写体の大きさに合わせて精度よく分割する必要があるので、小さい領域分割結果を基にパラメーターを調整し、順次、被写体に合った大きな領域に統合できる手法を考案し、その基本性能を確認した（図2）。

図2　小さい領域分割結果に基づく指定領域抽出の例

■メタデータ制作フレームワーク

メディア解析を利用した効率的なメタデータ制作環境として提案しているメタデータ制作フレームワーク（MPF:Metadata Production Framework）のMPF Ver.2で採用しているメタデータモデルを、EBUの活動を通して拡張し、ISO/IECにおいてMPEG-7の新しいプロファイルとする追補案を最終投票段階に進めた。また、この新しいプロファイルを採用したMPF Ver.3のリリースに向けて関連するリファレンスソフトウエアの開発を進めた。

■番組検索・推薦技術と視聴状況取得技術

視聴者の状況に応じてさまざまな番組の検索・推薦サービスを提供するユーザーインターフェースの実現を目指して、番組の検索・推薦技術と視聴者の視聴状況取得技術に関する研究を進めた。

番組の検索・推薦技術の研究では、NHKオンデマンド、番組公開ライブラリーの番組推薦支援システムで収集した操作履歴情報を利用して推薦順位を最適化する技術を開発した。また、映像説明文の類似度の計算手法と画像の類似度の計算手法を組み合わせた関連クリップ検索システムを、NHKクリエイティブライブラリー用の実験サービスとして公開し、利用者の操作性の評価や履歴の収集を進めた。

視聴状況取得技術の研究では、視聴中の番組やシーンに対する視聴者の興味度を推定する手法を検討した。興味度を推定するために、カメラによって観測された視聴者の行動などの情報とリモコンなどの操作情報を統合し、時系列ラベリング問題として機械学習を行う方法を考案した。評価実験の結果、約70％の正解率で興味を持って視聴している状態を推定することができた[3]。また、興味度の推定に寄与する情報として、番組映像への注意の向け方（注目度）の指標となる情報をカメラ映像から取得する手法の開発を進めた。注目度の指標となる情報として、2011年度は、顔の向き、顔の表情の変化度、身体動作の検討を継続したほか、新たに瞬目間隔、視線変動量の取得法の開発を進め、瞬目間隔・視線変動量・身体動作の3種類の情報を用いて番組に対する注目度を推定するシステムを試作・評価した。また、顔の表情の変化度の検出では、人ごとに顔登録が必要であった処理を多くの顔表情画像を事前学習し、顔を登録する必要のない手法に改善した。

興味度推定技術の応用として、推定した興味度に応じて推薦番組を変えるユーザーインターフェースUTAN（User Technology Assisted Navigation）を試作し（図3）、視聴者プロファイルの推定手法などの開発に必要な学習データを実際のテレビ視聴の環境で収集した。

また、人物行動認識手法の基礎検討として、人混みの映像を対象として、特定人物の行動から抽出した多数の軌跡特徴をクラスタリングして、人物単位での行動を検出する手法を考案し、国際会議での発表を通じて、その精度と今後の課題を確認した。

（注）本来の放送画面を縮小して角に寄せ、生じたL字型のスペースに他の情報を表示する画面

図3　UTANタブレット端末での操作風景

〔参考文献〕
(1) 後藤, 宮崎, 田中, 相澤: “更新履歴による注目度を利用した番組検索結果のリランキング,” FIT2011, D-025 (2011)
(2) 河合, 藤井: “エッジ共起を考慮した回転・スケール不変な局所画像特徴と一般物体認識への適用,” 信学技報, Vol.111, No.349, CS2011-78, IE2011-102, pp.139-144 (2011)
(3) 苗村, 高橋, 山内, 藤井: “CRFを用いたTVの興味視聴区間の推定手法,” 信学技報, Vol.111, No.430, PRMU2011-86, SP2011-101, pp.1-6 (2012)

3.6　視聴者心理状態推定技術

番組がどう見られ、どのような心理的効果をもたらすのかを分析するために、視聴者の脳活動、視線や視聴態度などから心理状態を推定する研究を進めている。

視線の動きの分析では、2011年度は、スーパーハイビジョンの見られ方を分析するために、広視野の映像を見ている人の視線の動きを非接触で測定する装置の開発を進めた[1]。被験者の瞳孔と角膜反射光の位置を検出して視線を測定する装置を複数台利用し、おのおのの測定結果を統合して測定範囲を拡張した。その結果、広視野映像を観視するときの自然な頭部動作を許した場合でも、測定誤差2度程度で水平方向80度の広視野測定が可能となった (図1)。

番組映像のどの部分に視聴者の視線が向けられるのかを決定する要因には、興味や意図に由来するトップダウン要因と、色や輝度といった画像の物理特徴に由来するボトムアップ要因があると言われている。2011年度は、このうちボトムアップ要因による視線移動を示すモデルとして提案されている顕著性マップに、画像中に複数存在する潜在的に視線を引付けうる要素が与える効果を追加した視線移動のモデルを提案した[2]。これにより視線が従来のモデルでは予測できない原因を分析できるようになり、広視野映像の効果の解析にもつながると考えられる。

図1　広視野・非接触視線測定システム

番組視聴中の視聴者の心理状態の変化を、同時に計測した脳活動データから解読することを目指した研究を行っている。2011年度は、対象としてお笑い番組を取り上げ、番組視聴時の脳活動データに多変量解析技術を適用し、視聴時に「おかしい」と感じているかどうかの判定を行った。その結果、側頭部にある「角回」と呼ばれる脳部位の反応を用いて、主観的な状態を高い精度で判定することができた。

映像による身体への悪影響を防止する技術の開発を進めている。2011年度は、映像酔いの原因となる動揺映像を自動的に修正するアルゴリズムを開発し、その効果を心理実験を行って検証した。

〔参考文献〕
(1) 小峯, 澤畠, 森田: “複数の眼球カメラを用いた視線測定範囲の広視野化,” FIT2011, J-039 (2011)
(2) 梅田, 小峯, 比留間: “ボトムアップ要因による視線分布の推移推定,” FIT2011, J-038 (2011)