

司会進行を支援するウェアラブルMCシステムの設計と実装

Design and Implementation of a Wearable MC System

岡田 智成 山本 哲也 寺田 努 塚本 昌彦*

Summary. どのようなイベントにおいても、司会進行のスムーズさはそのイベントの成否に大きな影響を与える。しかし、一般に司会をスムーズに行うためには、台本を綿密に覚える必要があり、さらに突発的に起こる様々なトラブルに対応する必要がある。また、司会進行ではイベント空間における「間」を適切に読んで場を取り仕切ることが必要であるが、熟練した司会者でないと間をうまく扱えない。そこで本研究では、ウェアラブルコンピューティング技術を活用し、これらの問題を解決するウェアラブル司会システムを提案する。提案システムを設計、実装し、2008年12月に行われた神戸ルミナリエのイベントにてプロトタイプシステムを実運用し、提案システムを用いることで司会進行をスムーズに行えることを確認し、実運用の際の問題点に対処するために音声による指示、台本のトラッキング機能、アドリブ検出機能などを実装し、より実運用に適したシステムが開発できた。

1 はじめに

司会者とは、学会や舞台、結婚式などの進行を司り、イベントを行う上で重要な役割を果たす人である。一般に、ステージにおける司会進行では、綿密な台本に基づいて訓練された司会者が進行を行う必要があり、あらかじめ設定された時間通りにイベントを進めなければならない。そのため、急に司会を依頼されてもすぐに対応できるものではなく、たとえすべて台本を覚えたとしても、本番で忘れてしまう可能性があり、発言を適切なタイミングで行うことも難しい。また、ステージ上での突発的なトラブルへの対処や聴衆からの質問への対処などさまざまな問題に臨機応変に対応する必要がある。しかし、経験の少ない司会者では、台詞や台本の流れを覚えるための準備に膨大な時間を要するだけでなく、聴衆の状態に配慮したスムーズな司会を行うことができない。

そこで、本研究では、ウェアラブルコンピューティングの特徴を活かして、ステージにおける司会サポートを行うウェアラブル司会システムを提案する。

ウェアラブルコンピューティングとは、常時コンピュータからのサポートを受けられ、頭部装着型ディスプレイ (HMD: Head Mounted Display) を用いることで他人に気づかれないで情報を取得できる [1]。また、装着型インタフェースを用いることでユーザーのジェスチャによる入力を行うこともできる。

提案システムでは、司会進行のために必要な情報を HMD に表示させることで、経験の少ない司会者であっても短い準備期間でスムーズな司会を行える

ことを目指す。

以下、2章では関連研究について説明し、3章でシステム設計と実装について述べる。4章で評価実験について考察し、5章で実運用について述べる。6章でまとめを行う。具体的にはイベントを進めるための必要な情報を表示させることで、本番で台本の流れを忘れてしまうことを防ぎ、落ち着いて司会できるだけでなく準備の時間を大幅に削減する。また、台本の見直しや司会評価のために、司会者が台本通りに話したかどうかを検出する機能を実装し、アドリブ箇所を取得できるようにする。

2 関連研究

これまで様々なプレゼンテーション支援システムが報告されている。2006年に栗原らによって発表されたプレゼン先生 [2] は、プレゼンテーションツールと連携しつつ、マイクおよび Web カメラから得られた発表者の音声および振る舞いを分析し、話速度、声の抑揚、聴衆とのアイコンタクトの度合いなどの指標をリアルタイムに発表者にフィードバックするプレゼンテーショントレーニングシステムである。発表終了後には、これらの指標をグラフとして可視化して提示することで、発表者のスキル向上のための事後解析を支援する。しかし、このシステムは、プレゼンテーションのトレーニングとして用いられるものであり、本番で使用することを目的としていない。本研究では実際の発表時に用いる司会支援システムを提案しており、設計コンセプトの点で異なる。また、プレゼンテーションは相手に情報をわかりやすく伝えることが目的であり、司会進行は、聴衆に指示をだし、会場の空気を制御することが目的であるため、支援すべき内容も異なる。

またウェアラブルコンピューティング技術を用い

Copyright is held by the author(s).

* Tomonari Okada, Tetsuya Yamamoto, Tsutomu Terada and Masahiko Tsukamoto, 神戸大学院工学研究科電気電子工学専攻

たステージを支援するプロジェクトが行われており HMD を用いた情報提示手法の有用性が確認されている。特に、本研究の先行研究として板生らによって行われたウェアラブル司会プロジェクト [3] は、本研究と同様にウェアラブル技術を用いた司会支援システムである。このシステムは単純にチャットシステムをウェアラブルコンピュータ上で動かして、オペレータからの指示を HMD 上で閲覧するものであり、また司会者は文字入力装置を備えていない場合が通常であるため、双方向のコミュニケーションが行えない。本研究は司会進行において支援すべき情報を整理し、それぞれに対して有効な支援機能を実装している点で異なる。

3 システムの設計と実装

3.1 司会支援装置

ステージにおける司会支援装置としては、プロンプタ (据え置きディスプレイ)、インターカムが代表的である。プロンプタやディスプレイは台本の変更やディレクターからの指示を表示させることができ柔軟性が高い。しかし、設置コストが高くステージの特定の場所に固定されてしまう。また、現場にいるスタッフはその画面を確認できない。これに対し、インターカムはハンズフリーで即時性の高い通信が可能であり、情報の共有が出来るが、情報量が少なく聞きなおしが出来ない。

このように現在の支援装置はスタジオ内での使用に特化しており、ステージ上で支援することを想定されていない。それらの問題点を解決するために提案システムはウェアラブルコンピューティング技術を用いる。まず、情報提示装置として HMD を使用する。HMD を用いることで情報提示装置を設置する必要がなくなり、ステージ上で様々な向き、場所で台本などの情報を得ることができる。さらに裏方のスタッフが HMD を装着することで舞台上の人だけでなく裏方のスタッフも情報を共有できる。HMD を用いる問題点として、画面の小ささと見た目の悪さが挙げられるが画面の小ささはシステムの工夫で改善される。HMD の不自然さについての実験もおこなわれており [4]、HMD を用いることは聴衆にとってそれほど不自然ではないことがわかっている。さらに、小型の入力インターフェースを用いることでオペレータへのフィードバックが可能になり、ステージ上のトラブルを瞬時に把握、対応できる。ステージ上のトラブルに対応する際、より即時性の高い対応が期待されるためイヤホン装着型インターカムと同様のシステムを実装している。

3.2 システム要件

ステージにおける司会進行では、綿密に台本を覚え、あらかじめ設定された時間通りにイベントを進

めなければならない。たとえば、台本をすべて覚えたとしても、本番で忘れてしまう可能性や、過度の緊張によってうまく話せなくなったりする可能性がある。またステージ上でのトラブルへの対処や聴衆からの質問への対処など様々な問題に臨機応変に対応する必要があり、司会者が困惑してしまった場合、司会者の不安感が聴衆に伝わり安心感のあるステージを進めることができない。さらに、司会者は演目と拍手の間を適切な時間に制御する役割をもっており、経験の少ない司会者は適切な間を作ることは難しい。そこで、本研究で提案する司会支援システムに求められる要件として下記の 3 点を挙げる。

1 点目、HMD に司会に必要な情報を表示する。台本、次のスライド、残り時間、オペレータからの指示、会場の状況といった司会進行の助けになる情報を HMD に表示する。また、HMD の画面は小さくて見にくいので、HMD から 1 度目を離すとどこを読んでいたかわからなくなる。そこで台詞を音声認識によって解析し、すでに読み終わった部分の台本の背景色をつけることで、目を離してもすぐに今読んでいる場所を確認できるようにする。また、司会者は台本をそのまま読み上げるだけでなく、台本に書いていないアドリブで会話を行う場合がある。よって台本とアドリブを区別して認識する必要がある。

2 点目、オペレータと交信し必要な情報を得る。オペレータは、司会者に指示をだす役割をもち、司会進行の助けになる情報を送信し、突発的なトラブルへの対応策を送る。具体的には、司会者の見える位置で台本や画像を送信したり、司会者にトラブルの対処法を送る。その時の、司会者側からの通信は、困惑しているのが観衆にばれないように秘匿性の高いインタフェースでやりとりし、コマンドを送ることによって司会者はオペレータと交信する。オペレータからの指示に気づかせるため、システムは指示を受けると司会者に音声で知らせる。また、緊急を要する際はオペレータから直接音声で指示を行える。

3 点目、適切な間で司会進行ができる。司会をする上で最も大切なのは「間の取り方」である [5]。間とは、狭義には「話さない時間」である。この時間を適切にとることで、観衆は心地よくイベントに参加できる。間は多くの経験を経て初めて制御できるものであるため、提案システムでは音響処理技術を用いて自動的に適切な間を司会者に知らせる。

3.3 システム構成

本研究で提案するシステムの構成を図 1 に示す。提案システムは、司会者とオペレータの 2 人で扱い、図で示している通り司会者は小型の PC、イヤホン、マイク、インタフェース、HMD を身につけ、オペレータはノート型の PC とマイクを用いる。司会者は HMD を装着することで、オペレータから送られてくる情報を閲覧する。また、司会者が台本を読ん

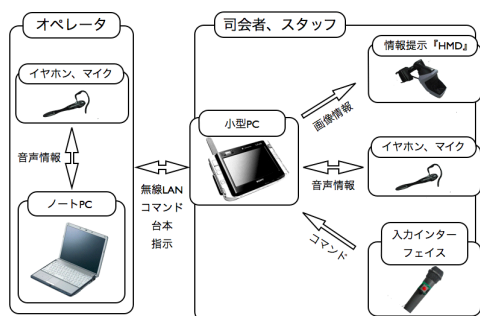


図 1. システム構成

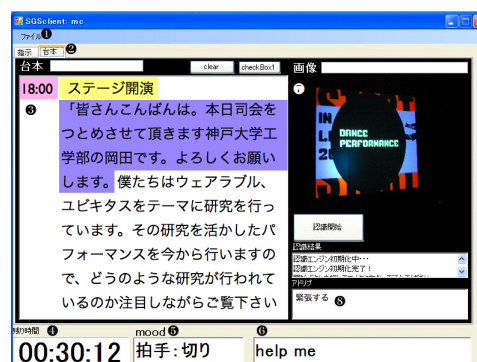


図 2. 台本画面

でいる際、HMD の画面から目をそらすことで文字の位置がわからなくなることを防ぐため、司会者が装着しているマイクから音声を取得し、DP マッチングを用いて認識された音声で台本かアドリブかを認識させ、台本ならばその部分の文字の背景色を変更させる。さらに、オペレータからの指示で重要かつ緊急を要する場合、音声チャットで直接指示できる他、文字の大きさ、色を変更できるようにし、指示をわかりやすくする。このようにすることで要件 1) を満たす。

オペレータは PC を操作し司会者に直すところや変更点などの指示を与える。2 台のパソコンは無線 LAN によって通信しており、双方離れていてもデータのやり取りが可能である。また、司会者から発信するために小型のジョイスティックをマイクの後ろに取り付けて操作を行う。こうすることで司会者の操作は聴衆に気づかれにくくなるだけでなく、細かい操作も可能となる。さらに、オペレータからの指示を司会者に気づかせるため、司会者側が指示を受信するとイヤホンからその指示の音声が再生される。こうすることで司会者は台本を読んでも指示に気づくことができるようになり、要件 2) を満たす。

また、オペレータ側の PC のマイクがステージの音を取得し、高速フーリエ変換 (FFT) することで観衆の声と拍手の違いを認識させる。そこで拍手のタイミングや場の雰囲気を読み取り司会者側に伝えることで、司会者は適切な間を作ることができ要件 3) を満たす。

3.4 システムの機能

構築したアプリケーションは、司会者用とオペレータ用に分けられる。以下、それぞれについて詳細に説明する。

まず司会者の HMD に表示される画面を説明する。司会者を行いながら小さな画面を見るため、フォントは大きく設定している。以下、アプリケーションの操作方法は図 2、図 3 を用いて説明する。HMD の 1 つの画面に全て情報を載せると文字が小さくなるの

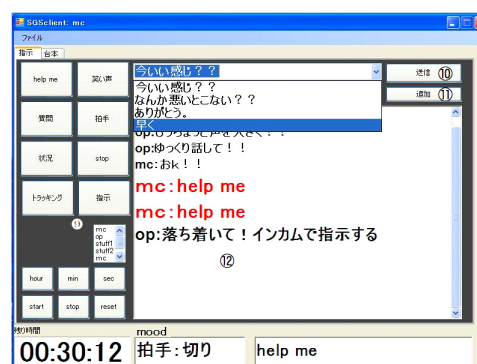


図 3. 指示画面

で、表示内容を台本画面、指示画面を選択する。図 2 は台本の表示画面であり、台本、画像、残り時間、間のタイミング、アドリブと認識された文字列を表示している。尚、台本は正規表現を用いて時間、台詞、イベントを区別できるようになっており読み終わった台詞の背景色が変化する。図 3 はオペレータとの通信画面であり、オペレータとの会話履歴を表示している。ヘルプボタン表 (??参照) で司会者からオペレータに指示を仰ぐことができる。また、司会者は事前に会話を登録することができ、オペレータに送信ボタンで送信できる。

次にオペレータが操作する画面について説明する。画面は司会者と同じような配置になっている。オペレータは主に司会者に指示を送る。この時、文字の大きさと色をコマンドによって変化させられる。コマンドは@の後にサイズや色などを指示の前に打ち込むことで文字を変化させられる。例えば、@big@red と打ち込めば文字を赤く大きくでき@beep と打ち込めば、司会者にメッセージの到着を音で知らせられる。そのほかに、台本、トラッキング位置の変更や画像の送信、会場の状態を把握できる。さらに、音声チャットで直接司会者に指示ができる。

表 1. ジョイスティックのコマンド割り当て

操作	動作
左に傾ける	マウス機能に切り替え
上に傾ける	コマンドスクロール
下に傾ける	コマンドスクロール (逆)
右に傾ける	画面の切り替え
クリック	決定
ダブルクリック	緊急コマンド

3.4.1 台本のトラッキング

上で述べた通り、ステージで司会を行いながら HMD に表示された文字を読むことは難しく、一度目をそらすとどこを読んでいたかわからなくなる。さらに、HMD の表示可能な解像度はそれほど高解像度でなく (800 × 600 ピクセル) 表示できる台本の範囲が限られている。よって、オペレータが台本の表示をスクロール、またページ送りを行うには作業回数が多くなる。そこで、音声認識器 Julius[6] を用い、読み終わったところの台本の背景色を変更し、スクロールを行う。しかし、ステージ上で話す音声を高い精度で認識することは難しく、司会者の台詞が長ければ認識率も下がる。一般のロボットの操作のような音声認識インタフェースでは、認識に失敗しても繰り返せるが、司会者の台詞を 2 回言うことはできないため、確実な認識が求められる。さらに、司会者が台本どおり台詞を読むとは限らずアドリブで会話をを行う場面も存在する。

そこで、音声認識した文字列が台本の文かどうかを認識させる。判定手法を図 4 に示す。まず台本、認識された語句をひらがなに変換し、認識された語句を 1 文字ずつスライドさせながら DP マッチングを行って、1 番低いスコアを抽出する。この時、認識される文字列は毎回語数が違うため正規化を行う必要がある。そこで 1 番低いスコアを認識された文字数で割り、正規化を行う。この正規化を行った値が閾値より低ければ台本と判断してトラッキングを行い、そうでなければアドリブと判断しタイムスタンプ付で保存する。保存され文字列はイベント終了後に確認でき、台本の変更や確認作業に役立たせることができる。万が一台本のトラッキングに失敗し、全く違うところに台本が移動した時のために、オペレータからのトラッキング位置の変更も可能にしている。

3.4.2 司会者のインタフェース

司会者は、小型の無線ジョイスティックを用いてシステムを操作する。このジョイスティックをマイクの後ろに取り付けることで操作しているのが気づかれにくくなる。また、用いるジョイスティックは方向選択だけでなくクリック機能もついており、無線

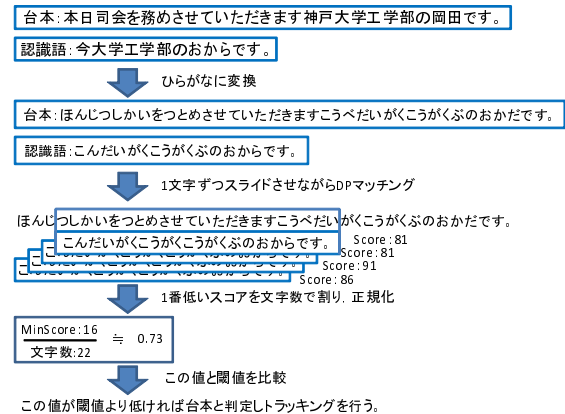


図 4. 台本のトラッキング判定手法

通信ができるため、デバイスと司会者が装着している PC との間にケーブルを必要としない特徴がある。このデバイスを用いて、司会者からコマンドを選択、決定しオペレータに指示を仰ぐ。コマンドの数はその場で増やすことができ、必要な言葉を入力することでコマンド数を増やすことができる。また、司会中はマウスを用いた細かい操作は司会者に操作負荷がおおきいので、表 1 に示すようにジョイスティックの傾ける方向でコマンドを割り振っている。

3.4.3 適切な間を指示する機能

演目の間を読むことをシステムによって支援する。上述した通り、オペレータ側の PC のマイクがステージの音を取得し、FFT によりマイクから入力された音を周波数ごとのスペクトルに分割し、スペクトルの特性から拍手と声を区別する。この区別を行うため、取得音を 12 段階の周波数領域に分割し、高周波数があれば拍手、それ以外の音は声と識別する。また会場の音量のデータの過去 1 秒の平均値を取得し、拍手が一定時間鳴りやまない時には司会者に注意を促す。この指示は司会進行をスムーズに行うための適切な間 [5] に基づいている。拍手の扱いは、ゲストに拍手のきっかけを与える「煽り」、拍手が鳴りやみそうなタイミングで次の演目に移る「待ち」、拍手の音が峠を越えた瞬間に話し始めて拍手をさえぎる「切り」の 3 種類であり、それぞれのタイミングでシステムが指示を与える。また、注目を集めるテクニックとして、観衆がざわざわしているなか話し始めても静まらない場合、5 秒ほど待ってから話し始めることで注目を得られる [5]。このタイミングを提案システムで制御する。判定のためのフローチャートを図 5 に示し、音量と時間、拍手、声の組み合わせで指示を決定している。指示内容は、拍手の場合音量が下がり始めれば「切」、5 秒以上鳴り止まなければ「待ち」の指示を与えている。

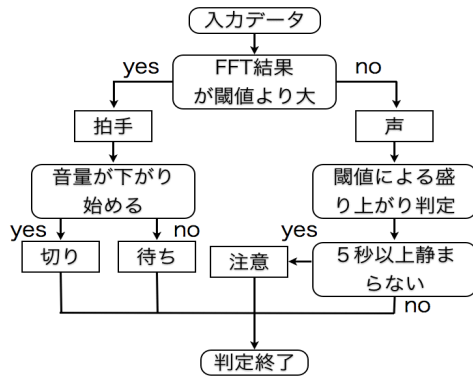


図 5. フローチャート

表 2. 判定結果

種類	正解	失敗	認識率
台本	56	11	84%
アドリブ	22	0	100%

4 評価実験

認識された文字列が台本かどうかの正誤判定に関する認識率の調査を行った。尚、台本はルミナリエで実運用した時、ステージ上で実際に用いた台本を用いており、アドリブもステージ上でパフォーマーとの掛け合いの台詞を用い、閾値は予備実験を行い決定した。実際に台本を読んで認識された文字列は67個(台本は40文)、アドリブを話して認識された文字列は22個である。その認識された文字列が台本かアドリブかを調査した結果を表2に示す。

この結果からこのトラッキング手法を用いることで、アドリブを100%認識することが出来、台本以外のことを話しても勝手にトラッキングが行われないう結果を得た。また、台本の認識結果は84%となった。この失敗の内訳は誤ってアドリブと認識された文字列が10個、台本の違う部分に認識された文字列が1個である。今回の認識率では5回に1回トラッキングされないことになるが、未トラッキングの次の文章でほぼ正しい台本位置に復帰できるので実用上問題はない。しかし、音声認識の結果が台本よりかなり違うかつ他の台本の部分にマッチングしてしまう場合は、オペレータがすぐさま変更する必要がある。

しかし、実際のステージではBGMや観衆の雑音が混入することで認識率が下がることが想定される。そこで、4通りの音環境化で実験を行った。その認識率を表3に示す。この結果から、4回目だけ全く正しい認識結果を得ることができなかったが、平均音量-25.86dBの環境では自分の声を確認するのも難しい環境であるため問題はない。その他の周辺音

表 3. 周辺音量がある状態での認識率

状態	認識率
平均音量-54.51dB, 最大音量-25.08dB	75%
平均音量-43.43dB, 最大音量-25.19dB	90%
平均音量-32.12dB, 最大音量 15.39dB	62.5%
平均音量-26.86dB, 最大音量-6.27dB	0%



図 6. ルミナリエステージでの実運用

量では認識率は多少下がったものの実運用上支障がでる結果ではなかった。さらに、司会者がステージ上で発言する際、BGMの音量は下げられると同時に観衆からの声援もほぼないと考えられるので、周辺音量が音声認識の障害にはならない。

5 実運用

5.1 ルミナリエステージでの実運用

提案システムの有効性を検証するために、2008年12月13日および14日に行われた神戸ルミナリエのイベントステージにおいてプロトタイプシステムを使用した。システムを利用している様子を図6に示す。ここで用いたシステムは提案システムの前段階のシステムであり、台本のトラッキングの機能と音声で指示を知らせてくれる機能を設けておらず、台本やオペレータからの指示を仰ぐ機能のみを持っている。準備期間としてイベントの前に5分ほど準備期間が設けられていたが、HMDなどを装着するだけなのでスムーズに準備を行うことができた。ステージでは、被験者である著者は緊張していたが、台本があったため落ち着いて司会を行うことができた。また、ステージ上のどこでもHMD上の画面を確認することができ、舞台上で場所を気にせず司会を行えた。ただし、1度HMDから目を離すと台本の位置がわからなかったり、オペレータの指示に気づかないなどの問題があり、これが提案システムに台本のトラッキングの機能と指示を音声で知らせてくれる機能を加えた理由である。またステージではテンキーを用いて操作を行ったが、装着している小型コンピュータのボタンに直接触れてしまい、フォー

カスを失って操作できなくなるといったトラブルがあった。このようなトラブルは、余計な緊張を生み出すと同時に、システムがなんらかの拍子に動作しなくなる可能性を示唆した。そこで、システムが動作しなくなった時のために、印刷した台本をもって、OS が落ちて動作できるマイコンを導入するなどの対策が必要であるとわかった。また実運用の際にはまだ空気読みシステムを搭載していなかったが、会場が屋外であると同時にBGMの音量が大変大きく、盛り上がりを認識することは難しそうであった。よって、BGMなどの音を除去し観衆の声援だけを認識する必要がある。今後はそれぞれの問題を解決し、司会を支援する様々なシステムを導入していきたい。

5.2 繰り返し運用

改善したシステムの検証として、所属研究室での論文紹介の司会もおこなっている。論文紹介では、発表者が英語の論文を聴衆の前でプレゼンし、質疑応答などの議論を行う。尚、発表者は4名、聴衆は20名前後で神戸大学の会議室で6回使用した。得られた感想は以下の通りである。

- 毎回台本を書く必要性がなく、準備期間が短くてすむ。
- 質問などが少ないとき自分一人の意見だけでなくオペレータと意見を共有できる。
- 台本の変更のようなトラブルも落ち着いて対処できる。
- 台本の場所を見失うことがない。
- 聴講しながら台本を確認できる。目を落として発表者の気にさせることがない。
- 時間通りに司会を行うことができる。

一方、下記のような問題も得られた。

- 会議室など移動がない空間ではHMDを用いる意味はあまりなく、PCの画面を用いるほうが有効である。
- 静かな場面ではオペレータから音声での支援は不可能である。
- メモなどの文字入力には現在のインターフェースでは不可能。

このようにウェアラブルコンピューティング技術を用いた司会支援システムが有効であると確認した。しかし、司会者の場所が固定である場合に限り、PCの画面を見ればよくHMDを用いる必要はない。スタッフ等に同じシステムを導入しより動的な支援を行うことも必要である。

6 まとめと今後の課題

本研究では、経験の少ない司会者がスムーズに司会を行うことのできるウェアラブル司会システムを開発した。提案システムは、オペレータと秘匿インタフェースを用いて交信することで、司会に必要な情報をHMDに表示させ進行を行うことができる。さらに音声認識をおこない読み終わった台本の背景色を変更させることで、1度目を離しても読んでいた位置がわかる。提案システムでは、単にHMDに情報を乗せるだけでなく、重要度に応じて指示の文字サイズ、文字色を変更でき、司会者が指示に気づくように音で知らせることもできるようになっている。また、司会者の能力で1番大事である「間」のとり方を知らせてくれるシステムを実装した。

今後の課題としては、より観衆に気づかれにくい入力インタフェースの提案が必要である。また、空気読みシステムにおいては、ステージのBGMによって盛り上がりの判別が難しくなることを考慮し、BGMの音を識別し除去する機能が必要である。さらに、システムがフリーズした時の対応策を提案する必要がある。また長時間司会を行うために電源をどのように確保するかという問題もある。一方、筆者以外の実践的運用も必要であり、それにより評価実験の客観性を増す必要がある。

アプリケーションの応用例としては、単に司会支援だけでなく、プレゼンテーション支援、舞台における台本支援、歌手の歌詞表示支援にも対応できると考えられる。

参考文献

- [1] 塚本昌彦: エンタテインメント用ウェアラブル・ユビキタスコンピューティング: エンタテインメントコンピューティングの事例, 情報通信学会技術研究報告, Vol. 44, No. 8, pp. 811-814, 2003.
- [2] 栗原一貴, 後藤真孝, 緒方 淳, 松坂要佐, 五十嵐健夫: プレゼン先生: 音声情報処理と画像情報処理を用いたプレゼンテーションのトレーニングシステム, 第14回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2006), pp. 59-64, 2006.
- [3] 板生知子, 塚本昌彦: ウェアラブル司会プロジェクト, 情報処理学会研究報告, pp. 5-12, 2003.
- [4] 池田 惇, 竹川佳成, 寺田 努, 塚本昌彦: 映像と連動したインタラクティブパフォーマンスのための演者支援システム, エンタテインメントコンピューティング 2009, pp. 1-8, 2009.
- [5] ゴトウライタ: 司会・幹事 段取りの仕方, 高橋書店, 2008.
- [6] 音声認識システムの開発・研究のためのオープンソースの高性能な汎用大語彙連続音声認識エンジン julius, <http://julius.sourceforge.jp/>.