

3.5 評価検証内容

(1) 技術的評価検証

a. ネットワーク機能

(a) メガカンファレンス技術

今年度メガカンファレンスを支援する MCU を用いた実証実験を行なうには至らなかったが、基礎研究の成果として、既存の MCU のボトルネックの特定や改善ポイントを特定することができた。また、今年度実施した多くの日韓交流イベントを通じて、本実証実験に参加する多くの日韓メンバーを確保することができたため、次年度以降メガカンファレンスを実施するための人的な環境が整った。

(b) マルチキャストIX技術

日本および韓国の研究ネットワーク間で MBGP を設定して、MBGP を用いた広域マルチキャスト経路制御のできる環境を整備したが、日韓いずれも MBGP 配下は単一の DVMRP あるいは PIM-SM によるマルチキャストネットワーク構成であったため、今年度はマルチキャストIX技術を十分検証できる広域な環境を手に入れることができず、詳細な検証は次年度以降の課題となった。ただ、日韓の基幹部分の MBGP 設定は完了しており、順次ネットワークを構成してゆけば本実証実験を行なえる環境を準備することができる。

(c) リクエストルーティング技術

今年度いくつかの VoD の実証実験を行ない、VoD 配送の基盤整備を行なった。

次年度はこの VoD をより広域に拡張し、「(a)メガカンファレンス技術」で記述したように、今年度の日韓交流で得られた多くの日韓メンバーの協力のもと、コンテンツ配送の際に負荷分散が必要な規模での実証実験を行なう予定である。なお、リクエストルーティングの基礎的な機能としては、DNS を利用した手法と、RTSP のリダイレクションを利用した手法によるものはすでに完成しており、次年度以降、今年度集めた VoD コンテンツ、視聴者を用いた大規模な実験を実施する予定である。

(d) 高信頼性マルチキャスト技術

今年度は、高信頼性マルチキャストのためのミドルウェアの実現、また、日韓交流の成果として韓国側の協力パートナーの確保、韓国パートナーによるオリジナルの高信頼性マルチキャスト手法の技術交流を行なった。「(a)メガカンファレンス技術」に記載したように次年度整備される日韓での安定した広域マルチキャストネットワークを利用して実証実験を行なう予定である。また、通信するデータとしては、「(c)リクエストルーティング技術」に記載した VoD コンテンツの巨大なアーカイブファイルがあり、この巨大なファイル群を各サロゲートに効率良く配送する実証実験を行なう。

b. インターフェース機能

インターフェース機能は、映像、日韓双方向翻訳機能を利用したチャット、音声、写真や3DCGなどの教材データなど、各種コンテンツを画面上に最適配置するとともに、講師が教材データへの書込みや、3DCGの回転操作等を行えば、参加者画面上においても、同様に表示されるデータ共有機能などを駆使して、コミュニケーションの円滑化や参加者の理解度促進を図るものである。

今回、日韓国際遠隔交流学习(中学校)、日韓国際遠隔講演(アート)、日韓企業間のビジネスマッチングなど、インターフェース機能を活用した各種アプリケーションを通じた実証においては、各アプリケーションを円滑に実施するため、当該アプリケーションの参加者との間で、事前に機能検証、改善(プロトタイピング)を行った。

(a) 参加者からの意見、要望 [プロトタイピング]

プロトタイピングにおける参加者からの意見、要望は、以下のとおりである。

表3 - 9 参加者の意見、要望

区 分	内 容
機能の単純化、 操作性の簡易化、向上	国際交流学习・遠隔講演における講師・講演者や、ビジネスマッチングにおけるメインの商談者は、当該アプリケーションの開催中、口頭による説明、解説交渉等に専念するケースが多く、各種の機能は必要最低限で、その操作も単純簡易にする必要がある。
	教材の選択が間違いなく行えるよう、ファイル名等ではなく、サムネイルの表示が有効である。
	教材への書き込みや、3DCGデータ等の回転等において、マウスによる操作は、ある程度の習熟がいるため、より簡易な操作方法を検討する必要がある。
機能の分離	自らチャット入力を行うことは通常困難であり、チャット入力については、機能を分離して補助者による入力が、効率的である。

(b) 機能改善

前項に示した参加者からの意見、要望に対応して、以下のとおり機能の改善を行った。

表 3 - 1 0 機能改善内容

区 分	内 容
機能の単純化、 操作性の簡易化、向上	操作者のミス等により、画面表示が消えたり、予期せぬ位置等に移動することのないよう、相手側映像、教材、チャットウィンドウ、教材選択アイコン等の画面枠を標準化した。 なお、相手側映像、教材画面については、表示画面サイズの異なる親画面、子画面を設け、簡単な操作で、教材画面等を親画面 子画面へ、また子画面 親画面へ切替できるものとした。
	教材選択アイコンはファイル名等でなく、教材のサムネイルが表示され、操作者が間違えることなく、認知できるようにした。また、今回、写真教材については、数百枚もあり、講師の状況に応じて必要な写真教材を選択できるよう、サムネイルアイコンを選択する方法に加えて、番号入力により、選択できる機能を付加した。
	教材への書き込みや、3DCG データ等の回転操作は、あたかも紙面上で行うように簡単に操作できる、液晶タブレットで対応した。
機能の分離	ユーザインターフェース画面と、チャット入力画面を分離して、相手側映像や教材等の画面を可能な限り、認知しやすく大きくするとともに、補助者よりチャット入力可能なよう対応した。

(c) 評価

これらの機能改善を行い、各アプリケーションの実証終了後における各参加者からの評価は、機能及び操作面に関して、3DCGにおける回転軸の色付け、教材データへの書き込みだけでなく、色を替えたり、内容を編集する等の、データ共有機能の向上について、改善要望があったが、これ以外では、特に支障なく、全般的には、操作しやすく、認知性もすぐれ、今回の実証実験以外での実教育やビジネス等でも十分に有効であることが確認できた。なお、今回、各アプリケーションを実施する前に詳細なシステム環境等について、各アプリケーション毎に作り込みを行ったものであるが、今後は各参加者においても、これら設定が簡単に行えるよう、汎用性の向上が必要である。

c . 日韓双方向翻訳機能

(a) 日韓双方向翻訳機能

日韓国際遠隔交流学習（中学校）、日韓医療スタッフ交流及び日韓国際遠隔講演（アート）において、日韓双方向翻訳機能を用いて行ったチャットのログデータをもとに、以下の方法により、有効性の評価を行った。

ア．評価方法

評価は、翻訳後の文章について、人が読んで意味が通じるかどうか、5段階で評価する主観的評価と、全文節数に対して正確に翻訳された文節数の比率をみる定量的評価により行った。

表3 - 1 1 日韓双方向翻訳機能の評価方法

区 分	評 価 方 法
評価対象	・各アプリケーションのチャットログ
主観的評価	・人が読んで意味が通じるかどうかを評価する5段階の主観評価 1 正確に翻訳されており、意味が完全に通じる 2 一部に誤訳があるが、意味はほとんど通じる 3 誤訳はあるが、意味は概ね通じる 4 誤訳が多く、一部しか意味が通じない 5 ほとんどが誤訳で、意味が全く通じない
定量的評価	・全文節に対して正確に翻訳された文節の割合を算定

イ．評価データ

表 3 - 1 2 日韓双方向翻訳機能の評価データ

主観的評価

	主 観 的 評 価					
	1	2	3	4	5	計
日韓医療スタッフ交流	150 (63%)	31 (13%)	25 (10%)	20 (8%)	13 (5%)	239 (100%)
日韓国際遠隔 交流学习	210 (67%)	33 (11%)	25 (8%)	26 (8%)	18 (6%)	312 (100%)
日韓国際遠隔講演 (ア-ト)	192 (74%)	34 (13%)	17 (7%)	10 (4%)	8 (3%)	261 (100%)
総 計	552 (68%)	98 (12%)	67 (8%)	56 (7%)	39 (5%)	812 (100%)

定量的評価

	翻訳精度
日韓医療スタッフ交流	86.9%
日韓国際遠隔交流学习	88.1%
日韓国際遠隔講演(ア-ト)	90.7%
総 計	88.6%

ウ．評価

評価データを総合的にみると、主観的評価では、“意味が完全に通じる”から“概ね意味が通じる”までを人が許容できる範囲とすると、約88%の文章がその範囲内であり、本ソフトウェアが有効に機能し、本ソフトウェアと組み合わせたチャットによる国際コミュニケーションが、意思疎通の円滑化の向上に寄与できたものと思われる。

また、定量的評価では、約89%の文節が正確に翻訳され、主観的評価の結果と近接しており、主観的評価の妥当性を示すものであると思われる。

誤訳の原因は次に示す様に、大まかに5つの項目に分類できる。

- A 翻訳辞書に登録されていない
- B 同音異義語を誤訳
- C 入力ミス(日/韓)
- D 動詞、形容詞などの変化失敗
- E その他

今回の翻訳結果の誤訳原因をみると、“翻訳辞書への未登録”と“語尾の変化の失敗”が誤訳の原因の半数以上を占めていることが分かった。これは、韓国語に多い同音異義語(例：“紹介”と“疎開”)や日本語にあって韓国語にはない、敬語表現などに起因するものがある。

同音異義語の翻訳、語尾変化の失敗については、システムの文章解析能力の改善を図る必要がある。また、翻訳辞書への未登録については、実証を重ねて、用途・分野に応じた専門用語のデータベースを充実し、日韓双方向翻訳機能の向上を図る必要がある。

今回の国際コミュニケーションにおいては、主観的評価において、12%程度、意味の通じない文章があったが、オンラインで繋がったチャットの利便性を活かして、同じ意味を別の表現や単語で入力し直し、意思疎通が図れるケースも多く、本実証実験では、有効に機能したものである。

表3 - 13 誤訳の原因分析

	原因評価					
	A	B	C	D	E	計
日韓医療スタッフ交流	42	18	14	45	0	119
	(35%)	(15%)	(12%)	(38%)	(0%)	(100%)
日韓国際遠隔 交流学习	37	23	9	27	7	103
	(36%)	(22%)	(9%)	(26%)	(7%)	(100%)
日韓国際遠隔講演 (アート)	33	18	1	22	0	74
	(45%)	(24%)	(1%)	(30%)	(0%)	(100%)
総計	112	59	24	94	7	296
	(38%)	(20%)	(8%)	(32%)	(2%)	(100%)

(b) 日本語音声認識機能

日本語音声認識機能については、個々の話者の発音等に応じて、認識精度が異なるとともに、一般的な日常会話と、大学講義等で話す専門的な内容とでは、用語が異なることから、これらによる認識精度の差異を考慮する必要がある。

また、本ソフトウェアが有する、話者の音響パターンの特徴や、単語辞書の登録学習機能の有効性についても考慮する必要がある。

これらのことを踏まえ、以下の方法により本ソフトウェアの有効性の評価を行った。

ア．評価方法

表 3 - 1 4 日本語音声認識機能の評価方法

区 分		評 価 方 法
評価対象者		・ 個々の話者の発音等による認識精度を評価するため、男性、女性複数の音声認識結果を評価 (男性：2名、女性：2名)
評価対象原稿		・ 日常会話及び、大学講義等における専門的な会話における認識精度を評価するため、新聞、技術論文について実施 (新聞：2原稿、技術論文：2原稿)
テスト方法	パターン1	・ 話者の音響パターンを学習せずに認識テストを実施
	パターン2	・ 話者の音響パターンを学習した後に認識テストを実施
	パターン3	・ パターン2による評価テストの結果、認識誤りのあった単語を辞書に登録学習(1回)させたうえで、再度、認識テストを実施
認識結果		・ 誤って認識した文節数 / 全文節数により評価

イ．評価データ

表 3 - 1 5 日本語音声認識機能の評価データ

評価方法 (パターン 1)

評価対象原稿 評価対象者	新聞 1 [177]	新聞 2 [264]	技術論文 1 [147]	技術論文 2 [155]	平均 (個別別)
女性 A	27	68	24	42	21.7%
	15.3%	25.8%	16.3%	27.1%	
女性 B	24	80	38	45	25.2%
	13.6%	30.3%	25.9%	29.0%	
男性 A	18	47	25	32	16.4%
	10.2%	17.8%	17.0%	20.6%	
男性 B	18	83	49	66	29.1%
	10.2%	31.4%	33.3%	42.6%	
平均(原稿別)	21.75	69.5	34	46.25	23.1%
	12.3%	26.3%	23.1%	29.8%	

注) []は総文節数、実数は誤認識文節数、ハ・セテージは誤認識文節数 / 総文節数、下表も同じ

評価方法 (パターン 2)

評価対象原稿 評価対象者	新聞 1 [177]	新聞 2 [264]	技術論文 1 [147]	技術論文 2 [155]	平均 (個別別)
女性 A	10	28	28	34	13.5%
	5.6%	10.6%	19.0%	21.9%	
女性 B	29	59	26	42	21.0%
	16.4%	22.3%	17.7%	27.1%	
男性 A	6	30	13	31	10.8%
	3.4%	11.4%	8.8%	20.0%	
男性 B	13	43	23	37	15.6%
	7.3%	16.3%	15.6%	23.9%	
平均(原稿別)	14.5	40	22.5	36	15.2%
	8.2%	15.2%	15.3%	23.2%	

評価方法 (パターン 3)

評価対象原稿 評価対象者	新聞 1 [177]	新聞 2 [264]	技術論文 1 [147]	技術論文 2 [155]	平均 (個別別)
女性 A	3	3	3	10	2.6%
	1.7%	1.1%	2.0%	6.5%	
女性 B	6	9	7	7	3.9%
	3.4%	3.4%	4.8%	4.5%	
男性 A	3	10	5	12	4.0%
	1.7%	3.8%	3.4%	7.7%	
男性 B	6	8	4	5	3.1%
	3.4%	3.0%	2.7%	3.2%	
平均(原稿別)	4.5	7.5	4.75	8.5	3.4%
	2.5%	2.8%	3.2%	5.5%	

ウ．評価

前項の評価データは、全体的に認識率は個体別、原稿別に格差が大きいこと、またパターン 3 で大きく認識率が改善されることが確認できた。

詳細にみると、個体別では、男性 A が評価対象の 4 原稿を通して、認識率が高い傾向が見られる。他方、女性 B は、同様に評価対象の 4 原稿を通して、認識率が低い傾向があり、個体の発声・活舌の特性に起因するものと考えられる。パターン 2 は、個体の音響パターンを学習した後の認識テスト結果であり、パターン 1 と比べ、総合で約 8% 認識率が改善されている。音響パターンの学習により、個体毎の発声・活舌の特性の差異による影響が、改善されることが分かる。

また、原稿別にみると、4 名の評価対象者を通して、技術用語や固有名詞の多い技術論文 2 の認識率が低いことが分かる。技術用語や固有名詞の多い会話においては、パターン 3 の単語辞書の登録学習機能が有効であり、認識率が大きく改善されている。

以上のことを考慮すると、日常会話においては、パターン 2 の結果のとおり、音響パターン学習を行うことにより、個体の発声・活舌による影響をある程度抑えることができ、また、技術用語や固有名詞の多い会話においては、パターン 3 の結果のとおり、単語辞書の登録学習機能が有効に機能していることが分かった。

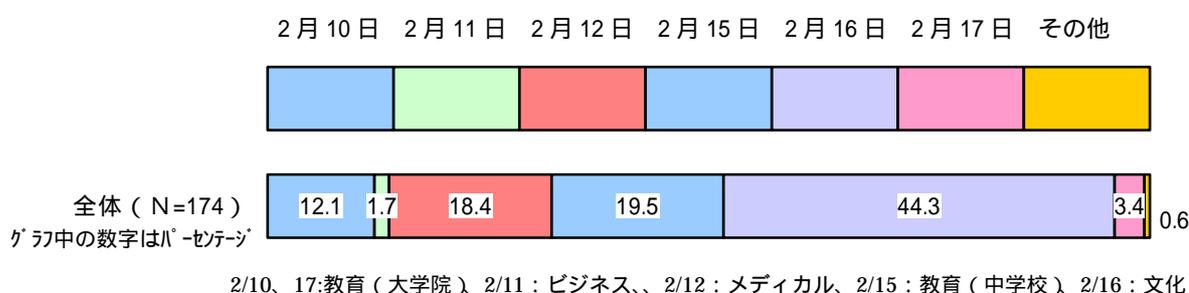
なお、同一の文章、同一の個体であっても、その時々で発声が異なり、また実際の会話では都度話す内容が異なり、逐次、新たな技術用語や固有名詞が出現することから、実証を重ねて、音響パターンや単語辞書のデータベースを充実し、日本語音声認識機能の向上を図る必要がある。

(3) 社会的評価検証

a . メガカンファレンスシステムの評価

メガカンファレンスシステム評価については、システムの利便性、システムの適用による業務効率化、映像及び音声の品質を評価項目とし、それぞれの項目に対して、参加者アンケートによる意識調査の結果に基づいて評価を行った。

なお、本システムは、日韓大学間での国際遠隔講義、中学生の国際遠隔交流学习、ビジネスマッチングなど、複数アプリケーションを通じて評価を行ったものであり、各アプリケーション毎の参加者数の比率は、グラフ3 - 1のとおりである。



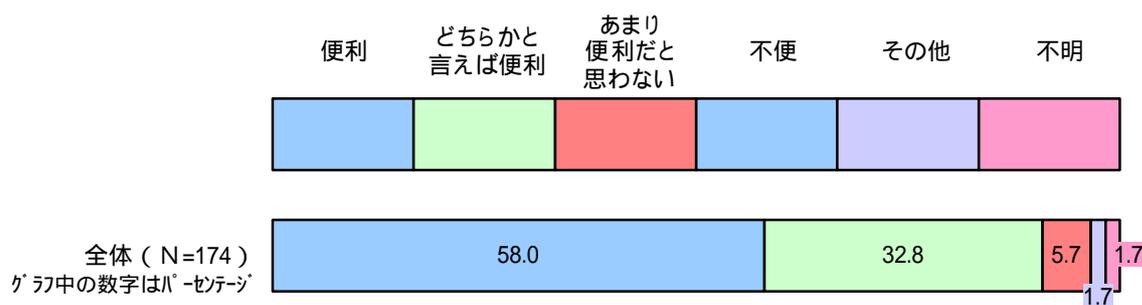
グラフ3 - 1 各アプリケーションの参加者の比率

(a) メガカンファレンスシステムの利便性

利便性については、本システムを、実際の大学の講義や企業等の商談などにおいて使用する場合、利便性を感じるか、活用されるか、という観点から、参加者の意識調査を行った。

その結果、約90%の人は便利だとの結果であり、実際の講義にも適用可能であるという意見が多く、利便性は十分高いとの評価を得た。また、2国間のみならず、世界中や、国内の他地域とも遠隔コミュニケーションを行いたいなど、より広い範囲での実証拡大の要望があった。

また、講師にとっては、高精細画像により受講者の表情や反応を確認できるため、それに応じて柔軟に進行を進める上でも、非常に有効であるとの意見もあっている。

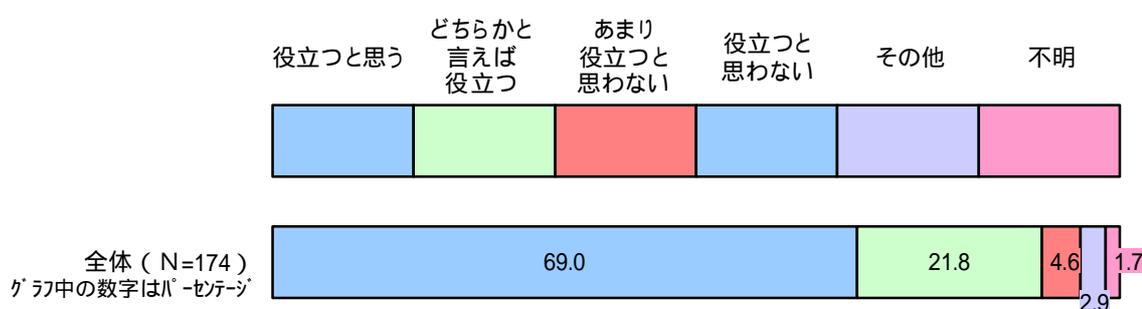


グラフ3 - 2 利便性に関するアンケート結果

(b) 講義・セミナー・商取引等における業務の効率化

現地に集合して実施する形態と比較して、経済的、時間的な効率化に有効であるかという観点から、参加者の意識調査を行った。その結果、約90%の人が業務効率化に役立つとの結果であり、今回実施した多人数による会議や講演などでも、高精細な画像と高度なインターフェース機能により、利用者は、本システムを利用すれば、現地に集合して、直接相手と対面する場合と同様な環境を体感でき、本システムが実用面でも十分有効であり、経済的、時間的な効率化に繋がると評価されたと考えられる。

また、特に変化の激しいビジネスや最先端技術に関する分野では、移動時間や、出張費用削減に加えて、タイムリー、スピーディーにコミュニケーションが行えるので、他者に先駆けてビジネスや研究を進める上でも、有効であるとの意見もあっている。



グラフ3 - 3 業務効率化に関するアンケート結果

(c) 映像及び音声の品質

映像や音声の品質について、参加者の意識調査を行った。

映像・音声の品質については、約70%の人が問題なし、残り約30%の人が問題ありとの結果となった。今回の評価は映像と音声を区別せずに、一体としてアンケート調査を行った結果であり、聞き取り調査を考慮すると、映像と音声では評価が異なる。

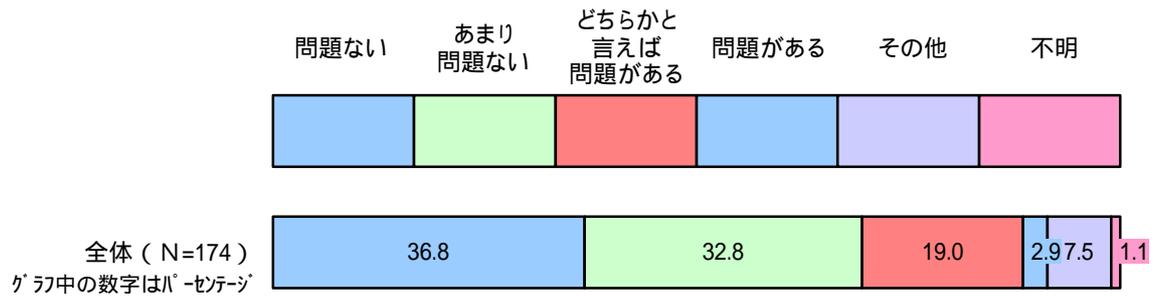
高精細の映像伝送が要求される、メディカル分野の内視鏡映像伝送や、文化分野の写真に関する遠隔講演の参加者に対する聞き取り調査では、映像品質については十分有効であるとのことで、高品質の映像が要求される用途にも、十分適用可能との評価を得ている。

一方、音声品質に関する聞き取り調査では、一部問題があるとの意見があっている。

具体的には、一部アプリケーションで相手側会場からエコーがあり、音声聞き取りにくかったこと、また、大会場で行ったアプリケーションでは、参加者の座席の位置により、音量が小さく、聞き取りにくい場合があるとの意見であった。

これらは、いずれも、本インターネット技術基盤自体に起因するものでなく、各会場にセッティングしたマイク、スピーカ等の周辺装置に起因するものと考えられる。

今後は、部屋の形状、建築材料、参加者の収容等を十分考慮した、音響システムを構成し、より効果のあるアプリケーション実証を行う必要がある。



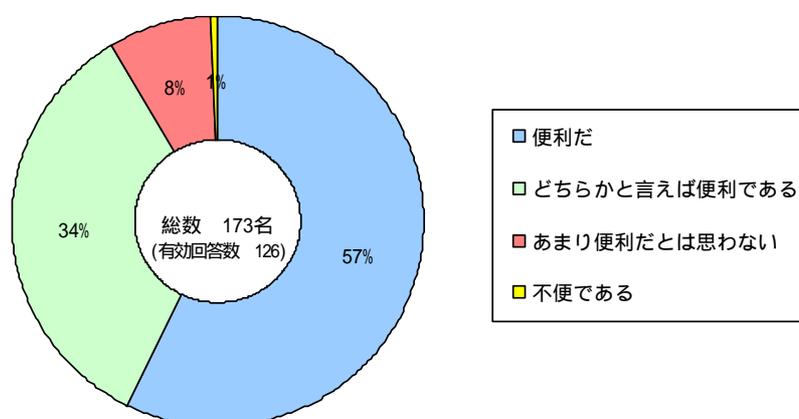
グラフ 3 - 4 映像・音声品質に関するアンケート結果

b . 韓国 TV 放送のリアルタイム配信の評価

韓国 TV 放送のリアルタイム配信の評価については、利便性、システムの適用による異文化理解促進への効果、映像及び音声の品質を評価項目とし、それぞれの項目に対して、インターネットモニターと福岡アジア美術館来館者に対するアンケート調査の結果に基づいて評価を行った。

(a) 利便性

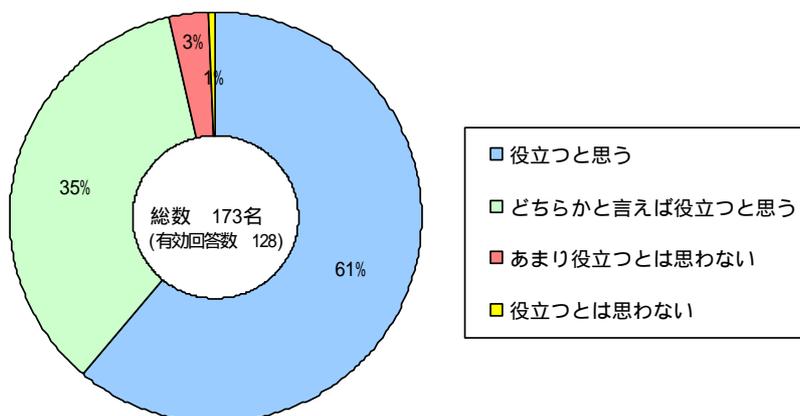
利便性について、約 91%の人が“便利だ”または“どちらかと言えば便利だ”との結果であり、本実証実験で実現したアプリケーションは、十分利便性が高いとの評価を得た。



グラフ 3 - 5 韓国 TV 放送のリアルタイム配信の利便性に関するアンケート結果

(b) 異文化の理解を深めるのに役立つか

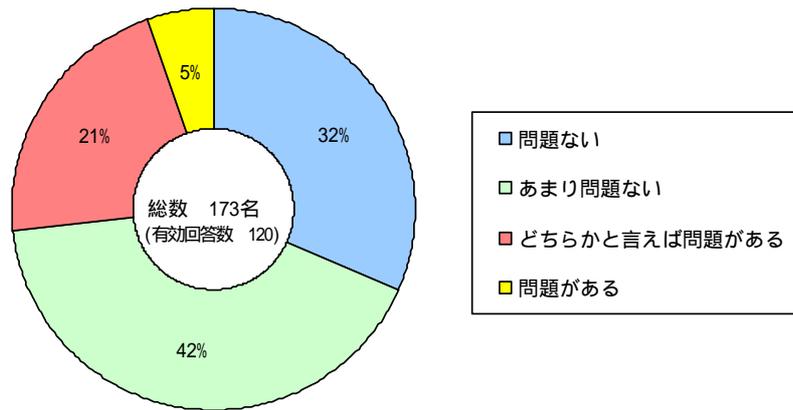
異文化への理解について、約 96%の人が、“役立つ”または“どちらかと言えば役立つ”との結果であり、本実証実験で実現したアプリケーションは、異文化への理解促進を図るうえで十分有効であるとの評価が得られた。



グラフ 3 - 6 韓国 TV 放送のリアルタイム配信の異文化への理解度向上への貢献度調査結果

(c) 映像及び音声の品質

映像・音声の品質については、約 73%の人は“問題ない”または“あまり問題ない”との結果であり、韓国 VOD 映像コンテンツの映像・音声の品質に対する結果に比べ約 12%低い評価となっている。これは、韓国 TV 放送のリアルタイム配信システムでは、PAL/NTSC の信号変換を行っており、韓国 VOD 映像コンテンツに比べ、若干映像品質が劣化したため、韓国 VOD コンテンツより低い評価になったものと推測される。



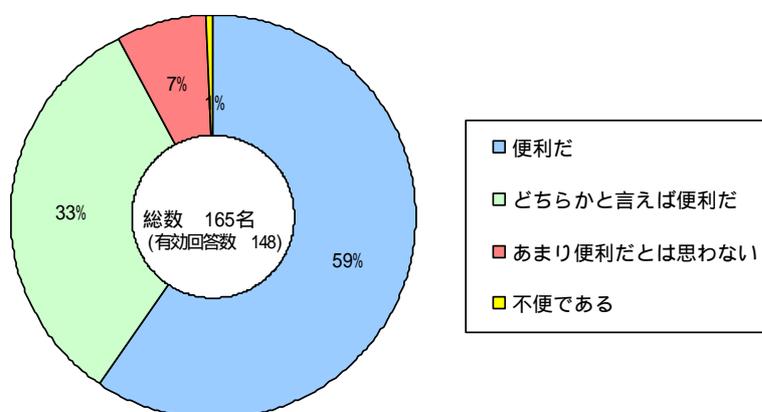
グラフ 3 - 7 韓国 TV 放送リアルタイム配信の映像・音声の品質に関するアンケート調査

c . 韓国 VOD 映像配信についての評価

韓国 VOD 映像配信の評価については、利便性、システムの適用による異文化理解促進への効果、映像、音声の品質及び操作性を評価項目とし、それぞれの項目に対して、インターネットモニターと福岡アジア美術館来館者に対するアンケート調査の結果に基づいて評価を行った。

(a) 利便性

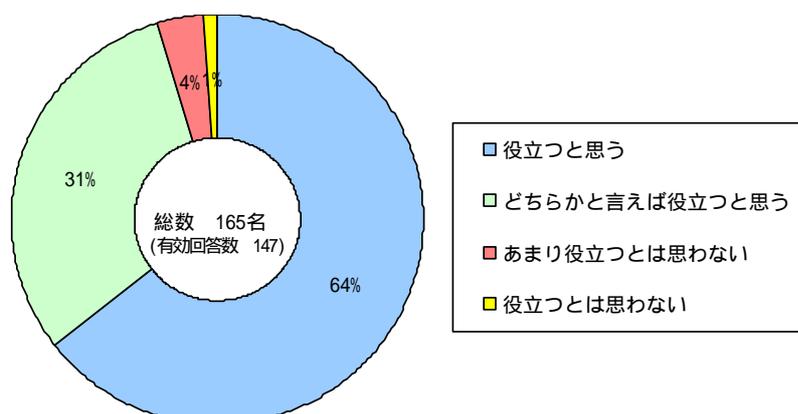
利便性について、約 92%の人が“便利だ”または“どちらかと言えば便利だ”との結果であり、本実証実験で実現したアプリケーションは、十分に利便性が高いとの評価を得た。



グラフ 3 - 8 韓国 VOD 映像配信の利便性に関するアンケート調査

(b) 異文化の理解を深めるのに役立つか

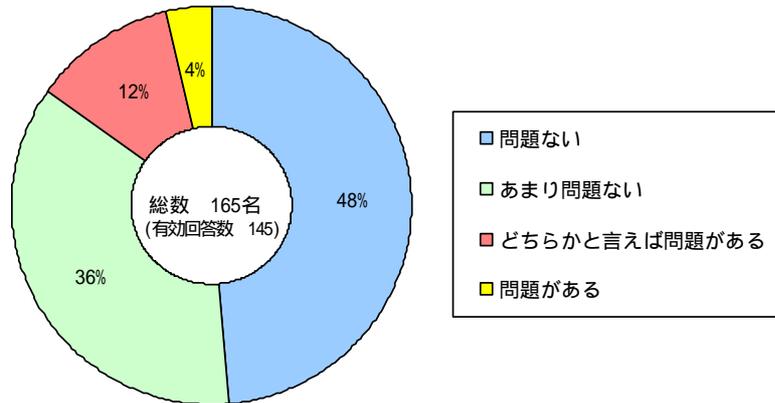
異文化への理解について、約 96%の人が、“役立つ”または“どちらかと言えば役立つ”との結果であり、本実証実験で実現したアプリケーションは異文化の理解促進を図るうえで、十分有効であるとの評価が得られた。なお、韓国の伝統文化に関するコンテンツのみでなく、現在の韓国を紹介する VOD コンテンツも紹介して貰いたいとの意見があった。



グラフ 3 - 9 韓国 VOD 映像配信の異文化理解度の向上への貢献度に関する調査結果

(c) 映像及び音声の品質

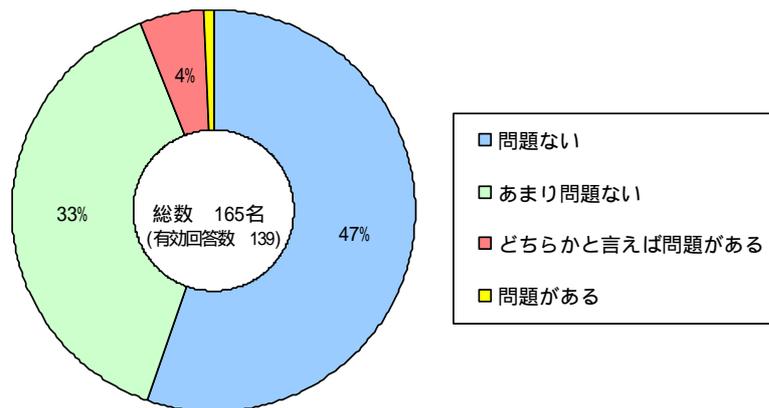
映像・音声の品質については、約84%の人が“問題ない”または“あまり問題ない”との結果であり、本アプリケーションの映像・音声の品質は、概ね高いとの評価が得られた。



グラフ3 - 10 韓国 VOD 映像配信の映像・音声品質に関するアンケート調査結果

(d) 操作性

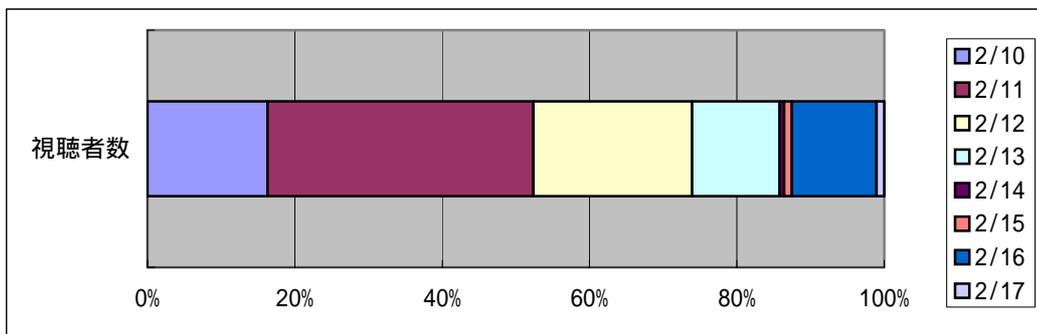
操作性については、約80%の人が“問題ない”または“あまり問題ない”との結果であり、本アプリケーションの操作性は、概ね良好との評価が得られた。



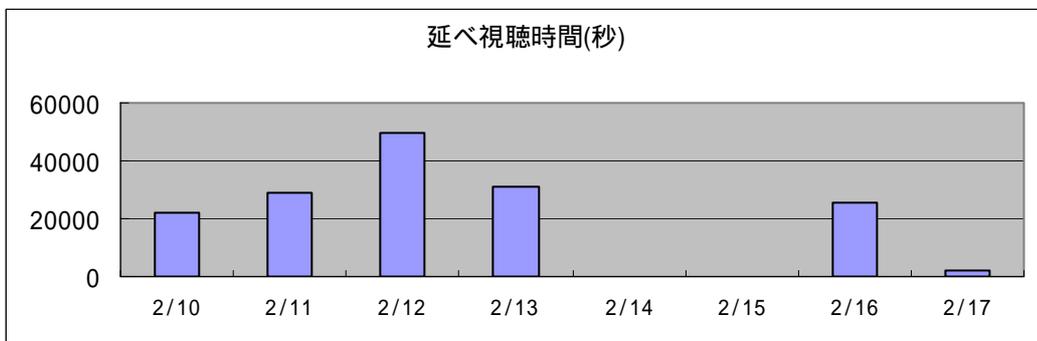
グラフ3 - 11 韓国 VOD 映像配信の操作性に関するアンケート調査

d. インターネットモニター視聴状況

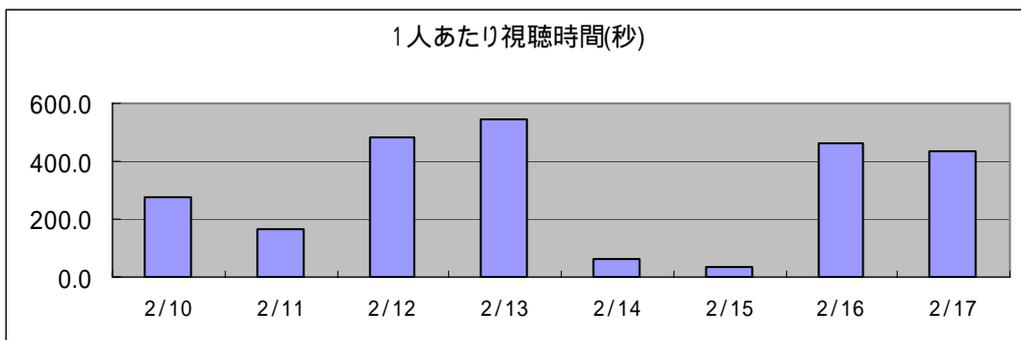
インターネットモニターの視聴状況について、グラフ 3-12 に日別の視聴者数を、グラフ 3-13 に日別の延べ視聴時間を、グラフ 3-14 に日別の一人あたりの視聴時間を示す。容易に推測されるように、平日より休日の方がアクセスが多く、またモニター配信実験の期間始めと、期間終わりでアクセスが多くなっている。より多くのモニター参加を促し、本プロジェクトのアピールに資するとともに、実証実験結果の評価精度の向上を図るため、実験期間の期中等において、適宜、参加を喚起するイベント等を実施することが有効であると考えられる。



グラフ 3 - 1 2 インターネットモニター視聴状況 (日別視聴者)



グラフ 3 - 1 3 インターネットモニター視聴状況 (延べ視聴時間)



グラフ 3 - 1 4 インターネットモニター視聴状況 (一人あたりの視聴時間)