

氏 名	ゴ 吳	ジュン 俊
学位(専攻分野)	博 士 (学 術)	
学 位 記 番 号	博 甲 第 150 号	
学位授与の日付	平成 9 年 3 月 25 日	
学位授与の要件	学位規則第 4 条第 1 項該当	
研 究 科・専 攻	工芸科学研究科 情報・生産科学専攻	
学 位 論 文 題 目	複数音声識別法を用いたデジタル映像の品質評価	
審 査 委 員	(主査) 教 授 田 村 博 教 授 西 村 武 教 授 黒 川 隆 夫	

論 文 内 容 の 要 旨

音声を識別する際に、発話者の映像を示すことによって識別の正答率を向上することができる。このことを話者映像効果と呼ぶ。本研究は話者映像効果を比較することにより、デジタル映像の品質評価を行った研究成果をまとめたものであり、話者映像効果を左右する幾つかの因子を見出すことに成功している。

話者映像効果の評価法として、混合音声を用いた複数音声識別法が、デジタル圧縮・再生映像の評価に有効であることを示した。圧縮・再生映像としては、低速 ISDN 回線を用いた ISDN 映像、MJPEG を用いた MJPEG 映像についての実験、更にデータ圧縮のモデルとしてストップモーション映像について実験を行っている。また中国人留学生の日本語に対する話者映像効果を測定し、滞在日数と話者映像効果との間の興味深い関係を示すと共に、清音と濁音の識別のように映像的に必ずしも明確な識別手がかりがないにも拘わらず、話者映像の効果が見られることを示した。

導かれた研究結果を要約すれば、次の通りである。

- (1) 静止映像には話者映像効果が認められないことを示した。ストップモーション映像では通常映像のほぼ半分程度の話者映像効果を認めた。ただし、唇音において、ストップモーション映像と通常映像の話者映像効果はほぼ同じレベルであった。
- (2) 通常映像と ISDN 映像の映像効果を比較した結果、ISDN 映像では映像効果が有意に低いことが確認された。映像効果の低下は唇音において著しい。話者映像効果を構成するものとして静的効果と動的効果を見出している。ISDN 映像における話者映像効果の低下は動的効果の減少に起因するとみなされる。
- (3) ISDN 映像では、映像効果が認められない被験者が30%に達する。現状の ISDN 映像 (64 kbit/s) の持つ重大な欠点と指摘される。これも静的効果と動的効果のアンバランスによるものと説明されている。
- (4) MJPEG の映像では、実時間再生が保証されているので、画質が多少劣化しても、動きが検知できる程度であれば、動的効果は十分に発揮される。それに対して画質は圧縮率と共に劣化するので、静的効果に依拠する非唇音の話者映像効果が低下する。
- (5) 以上のように同じデジタル画像であっても圧縮・再生の方法によって、話者映像効果に特性上の差のあることを示した。
- (6) 正答率と音声対の波形の類似性について検討を行った結果として、複数音声識別実験法に用いる音声単語の選択および組合せに必要な指針を導いた。これらは今後映像と音声の相互作用を説明する有効な基礎となる。
- (7) 中国人留学生の日本語単音節知覚に対する話者映像の提示効果を検討した。清音・濁音の識別に対す

- る話者映像効果は、動的効果、静的効果とも異なる第3の効果の存在を示唆するものである。
- (8) 複数音声識別実験から示唆される視聴覚相互作用のモデルを示した。

論文審査の結果の要旨

言語識別における視覚刺激の効果については、口話、読唇などと呼ばれ関心がもたれてきたが、健常者においてどの程度の重要性を持っているかを知ることは、たとえばマルチメディアの映像表示の価値を評価する上で大切である。また圧縮・再生映像の静的・動的画像と視覚効果の関係を明らかにすることも大切である。ISDN映像では話者映像効果が通常映像に比べて有意に低いことはある意味で当然ではあるが、話者映像効果に個人差が大きく、人によっては話者映像効果が無または逆効果であることを示した最初の結果として評価できる。さらにいくつかの圧縮・再生映像の話者映像効果を比較し、話者映像効果にいくつかの構成因子が存在することを明らかにしたのも本研究の成果と評価される。

人が人工的な環境と交わることが多くなると共に日常生活では目立たない機能障害に注意することが求められるようになってきた。データの圧縮・再生は将来のデジタル通信やマルチメディアの普及において重要な役割を果たすものであり、これらの技術に於いて、新しい形の情報障害者を生み出さないように細心の注意を払うことが求められる。そのためには情報機器の生理・心理・行動工学的評価法の研究が大切であると考えられる。本研究の結果は圧縮・再生技術を用いたデジタル映像についての、新しい評価法を確立する道を開いたものである。

本研究では話者映像効果には少なくとも2種類があり、静的効果と動的効果と呼んでいる。静的効果とは比較的ゆっくりと変化する映像の効果であり、母音や非唇音の知覚に効果を持つものと推定される。一方動的効果とは破裂唇音のような急速な映像の変化によってもたらされる効果である。通常映像では映像効果の個人差は必ずしも目立たないが、ISDN映像では静的効果に顕著に反応する視聴者や動的に効果に反応しない視聴者があり、全体として話者映像効果の低下に結びついていることが示された。

この理解と裏付ける意味で、ストップモーション映像、MJPEG映像の話者映像効果が検討されており、ストップモーション映像、MJPEG映像においても、高速効果は維持されること、MJPEG映像において圧縮率を高めると静的効果が劣化することなどが示している。

これらの結果はマルチメディア技術に不可欠なデジタル圧縮映像の機能評価のための基礎となる、独創的な研究であり、博士（学術）の学位授与に値する。