

研究概要報告書

資料-8

(1/2)

研究題名	音声に含まれている感情表現の分析とモデル化	報告書作成者	新美康永
研究従事者	新美康永 小林 豊		
研究目的	<p>本研究の目的は、音声に含まれている感情を察知したり、感情のこもった音声を合成するための手がかりを探ることである。音声を聞いたときに感じる話者の態度や感情は、話者の話し方と話されている内容とによってもたらされる。話し方は、発声の速さ、声の大きさ、声の高さなどの音響的パラメタを変えることによってコントロールされている。本研究では、この音響的パラメタと話されている内容とが、音声から聴き取られる話者の態度や感情とどの様に関係しているかを分析しモデル化する。話者の態度や感情を聴き取ることができれば、言語的内容以上の情報の伝達が可能となり、例えば、人間らしい対話システムの構築が期待される。また、音声合成の分野では、感情を合成音声に付与する研究は殆ど行われていないが、近い将来、視覚障害者に対する朗読サービスの自動化など、人間らしい感情のこもった音声を合成する技術への要求は高まってくるだろう。本研究はその先駆けとなるものである。</p>		

様式-9

研究内容

研究の内容 本研究では、音声を聴いたときに感じる感情が、話者の話し方や話の内容とどの様な関係にあるかを重回帰分析によって分析する。まず、多くの文を感情を込めて声優に読んでもらい、感情のこもった音声の資料をつくる。次に多数の被験者にこの音声を聴かせ、その印象から話者の感情を主観的に判断させ、感情の種類とその大きさの程度を5段階評価させることにより数量化する。この値をYとする。同様に音声の内容に相当する文を被験者に示し、それを黙読したときに感ずる感情の種類と大きさの程度を5段階評価させることにより数量化する。この値を X_1, X_2, \dots, X_n とする。一方、収集した音声を音響的に分析し、発声速度、基本周波数（声の高さに相当する）や振幅（声の大きさに相当する）の時間的な変化を求める。さらに、藤崎らによって提案されている基本周波数軌跡の生成モデルを用いて、段落の声立てやアクセントの大きさに相当するパラメタを抽出する。これらの音響パラメタを A_1, A_2, \dots, A_m とする。ここで、Y, X, Aは平均値が0、分散が1になるよう正規化しておく。Yが下に示すような線形結合によって表されるものと仮定し、その係数を最小自乗法によって求める。（これを重回帰分析という）

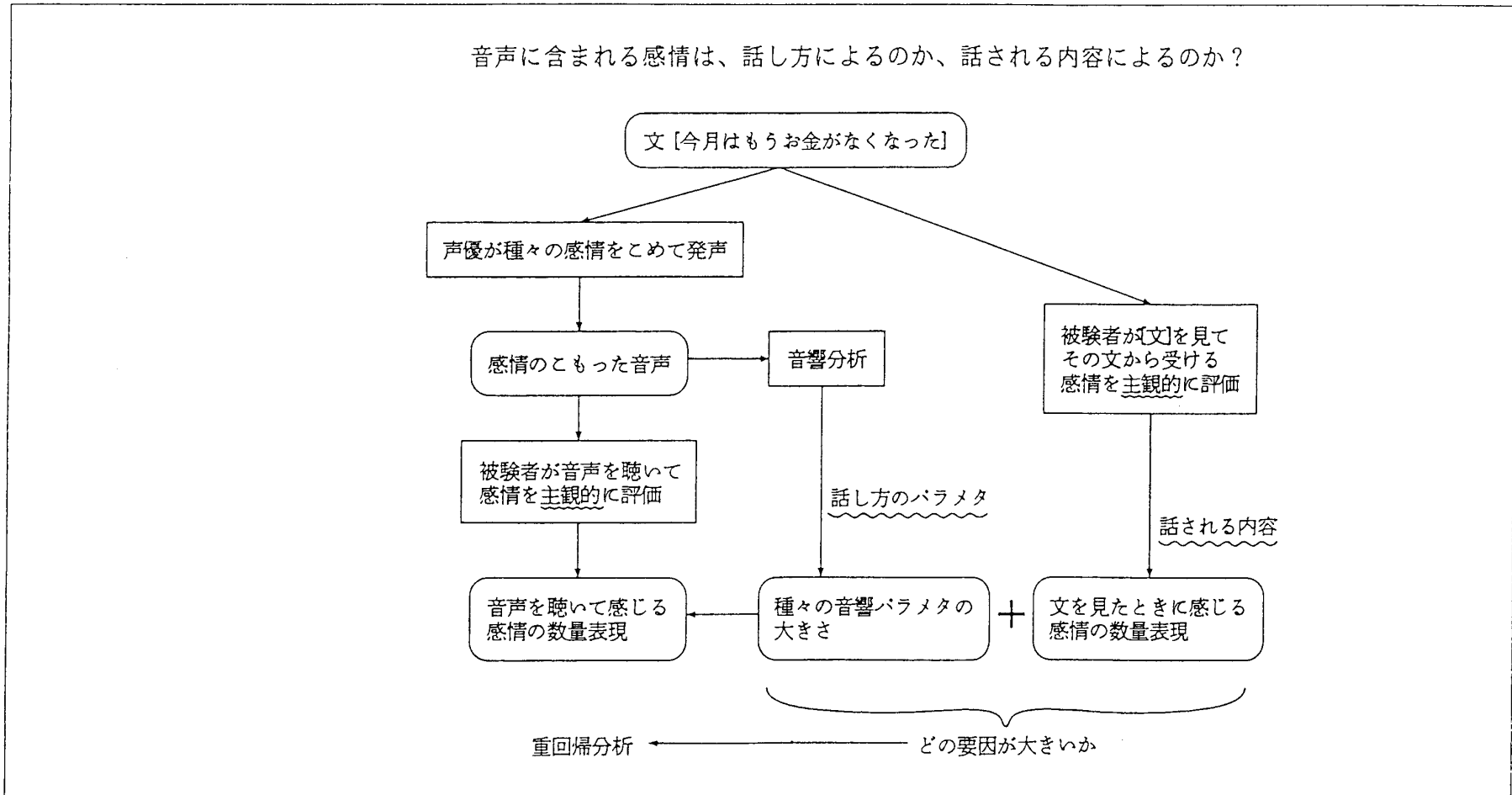
$$Y = a_1 X_1 + \dots + a_n X_n + b_1 A_1 + \dots + b_m A_m$$

これらの係数のうち値が大きくなったものがYに対する寄与率が大きな要因であったと判断する。

研究結果の概要 テキストから得た感情の大きさ（喜び、悲しみ、怒り、驚き）を X_i ($i=1\sim4$)、発話速度（モーラ/秒）、平均振幅、平均基本周波数（Hz）、基本周波数の標準偏差（Hz）を A_j ($j=1\sim4$)として、重回帰分析を行った結果、下表のような結果を得た。これから、①驚き以外では、発声内容のもつ言語情報が、聞き取られる感情に大きな影響をもつ。②各感情の大きさには音声の平均振幅が大きく影響する。③その他、声が低いほど怒りの感情を感じやすいこと、驚きでは発話速度の影響が大きいことなどが分かった。

表 重回帰分析の結果

	a_1	a_2	a_3	a_4	b_1	b_2	b_3	b_4
喜び	0.29	0.05	0.10	0.12	-0.11	0.27	0.11	-0.05
悲しみ	0.04	0.21	0.08	0.06	0.02	-0.21	-0.04	0.04
怒り	0.10	0.08	0.12	0.10	-0.02	0.53	-0.23	0.06
驚き	0.07	0.06	0.10	0.12	-0.32	0.48	-0.02	-0.00



(注： フローチャート図，ブロック図，構成図，写真，データ表，グラフ等 研究内容の補足説明に御使用下さい)

様式-10