

顔面特徴の選択的統合と表情および その強さの知覚に関するモデル¹

鵜 沼 秀 行*・長谷川 桐**

Selective Integration of Facial Features and Models of Perception of Facial Expressions and their Intensities

Hideyuki UNUMA and Hisa HASEGAWA

Abstract

Studies of perception of facial expressions have left unresolved issues of information processing models, including how visual features are extracted from the facial surface and integrated. In this article, information processing models of facial affect perception and empirical findings are reviewed with regard to the ability to predict the perceived facial expressions by human observers. Moreover, the authors propose a new model that indicates selective integration of facial features to perceive facial expressions and their intensities, and they summarize the empirical evidence for the model.

Key Words: face perception, facial expressions, facial features, selective integration

これまでさまざまな研究において、人間が他者の顔面にあらわれる表情をきわめて鋭敏かつ正確に知覚することが指摘されてきた (Ekman & Friesen, 1975; Ekman, Friesen, & Ellsworth, 1972)。特に, Duchenne (1862/1990) と Darwin (1872) の先駆的な研究以降, 多くの研究は表情の知覚と顔面の目や口の動きなどの要素的特徴の間の関係に焦点をあててきた (Ekman & Friesen, 1978; Essa & Pentland, 1994; Ellison & Massaro, 1997)。しかし, 代表的な研究としてあげられる Ekman ら (Ekman & Friesen, 1978) の表情記述システムにみられるように, 従来

*教授 知覚・認知心理学

**青山学院大学大学院文学研究科 博士後期課程

の研究は必ずしも顔面の視覚的特徴がどのような視覚情報処理を経て表情の知覚を誘発するのかを明らかにしてはいない (Sayette, Cohn, Wertz, Perrott, & Dominic, 2001)。そこで本研究は、視覚的特徴から表情が知覚される情報処理過程についてのいくつかのモデルを比較検討し、特に特徴の統合による表情知覚モデルの具体的問題について考察する。

表情知覚を説明しようとするモデルは、大きく全体モデル (holistic models) と特徴体制化モデル (feature configural models) にわけることができる (Ellison & Massaro, 1997; Schwaninger, Wallraven, Cunningham, & Chiller-Glaus, 2006)。全体モデルは、顔が部分に分割されない全体として表現されていることを仮定し、これは表情知覚に限らず顔の同定一般において前提とされる (Tanaka & Farah, 1991, 1993)。その根拠は、顔の部分特徴が単独で提示されたときよりも全体のなかで提示された場合に容易に知覚される (Tanaka & Farah, 1993) という実験事実にある。この仮定は、パターン認知のモデルにおける鋳型照合モデル (template-matching model) の前提に類似する (Ellison & Massaro, 1997) ともいえよう。

これに対して特徴体制化モデルは要素的特徴 (feature element) と特徴間の体制化に関する情報 (configural information) を区別する。たとえば Bruce (1988) によれば、この体制化に関する情報とは目と口の距離のような顔特徴の空間的関係をさしている。特徴体制化モデルは、たとえば顔が倒立提示された事態では要素的な特徴の知覚は比較的影響を受けないのに対して、体制化に関する情報の処理が影響を受けることで顔の同定や表情の知覚が困難になると仮定しており、この仮説を支持する実験結果も報告されてきた (Searcy & Bartlett, 1996; Leder & Bruce, 1998; Collishaw & Hole, 2000)。これらの結果を全体モデルで説明することは困難であろう。

上述のような全体モデルと特徴体制化モデルを比較検討する際の具体的論点として、(1) 要素的特徴およびそれらの関係がどこまで表情知覚に影響を与えるのか、そして (2) 知覚される表情が連続的に変化するか、それともカテゴリカルすなわち不連続に変化するか、という二点をあげることができる (Ellison & Massaro, 1997)。特徴体制化モデルは特徴とそれらの関係が表情知覚を規定すると仮定するが、全体モデルは特徴および特徴要素間による影響を仮定してはいない。また、全体モデルは一般に知覚される表情が異なる表情カテゴリ間で不連続に変化するととも同一カテゴリ内ではその弁別精度が低く、表情知覚がカテゴリカルであることを主張する点でカテゴリカルモデル (categorical models) でもある (Etcoff & Magee, 1992)。これに対して特徴体制化モデルは必ずしも知覚される表情のカテゴリカルな性質を仮定せず、むしろ特徴の刺激次元の変化に対して連続的に知覚される表情も変化すると考える (Ellison & Massaro, 1997)。

本稿は、以下でこれら二つの観点から最近のいくつかのモデルを比較検討し、さらに単純な特徴体制化モデルの限界を指摘しつつ、新たな選択的特徴体制化モデル (selective feature configuration model) を提案する。この選択的特徴体制化モデル (Hasegawa & Unuma, 2008a, b) は特徴間の重みづけられた選択的処理とそれらの統合による体制化をモデル化したものであり、そこでは知覚される表情の連続的変化と要素的特徴間の関係情報によって表情知覚が規定されることが仮定される。さらにこれらの仮定を支持する実験的な根拠 (Hasegawa & Unuma, 2008a, b) があわせて簡潔に紹介される。最後に表情知覚を説明しようとするモデルの今後の方向が論じられる。

特徴体制化モデルとその評価

表情知覚を顔面の要素的特徴から説明しようとした最近のモデルとして、Huber & Lenz (1993) をあげることができる。彼らは必ずしも知覚される表情そのものを論じているわけではないが、用いられた材料はブルンスウィック (Brunswik) 型の図式顔であって、それらは作成された時点で意図されたように (Brunswik & Reiter, 1937) 表情の知覚を誘発するものであり、また人間を対象とした多くの表情知覚ないしカテゴリ知覚の研究で用いられてきた (Reed, 1972; Nosofsky, 1991) ものである。Huber & Lenz (1993) はハト (pigeon) の弁別行動を図式顔を構成する要素的特徴、たとえば顔面上での目や口の高さなどの線形結合で説明しようとした。これを線型加算モデル (linear additive model: LAM) とよぶことにする。彼らの実験結果はこのモデルを十分に支持するものであったが、特定の視覚的特徴と知覚される表情の関係が検討されることはなかった。

これに対して、Ellison & Massaro (1997) はコンピュータグラフィックスで現実の顔を模した材料を作成し、その眉と口元という部分的特徴を操作してこれらから観察者が知覚する怒りと幸福の程度を測定した。さらに視覚システムへの刺激入力から表情知覚に至る顔刺激の部分的特徴変化の処理過程について、具体的かつ総合的モデルが提案された。彼らのモデルは表情知覚に限らず、音声言語知覚などの多様な情報を統合する事態での人間の一般的なパターン認知のモデルとして位置づけられ Fuzzy-Logical Model of Perception: FLMP (Massaro & Friedman, 1990) とよばれた。Huber & Lenz (1993) のモデル (LAM) が特徴の値の線型加算を仮定したのに対して、Ellison & Massaro (1997) のモデル (FLMP) では特徴の値の積が反応を決定するとされた。観察者による弁別と評定実験の結果、FLMP が LAM など他のモデルよりも観察者の反応を充分説明すること、また全体モデルではこのような要素的特徴の影響

を説明できないことが主張された。

Ellison & Massaro (1997) における FLMP の表情知覚への適用には二つの問題点を指摘することができる (Hasegawa & Unuma, 2008a)。彼らの実験で用いられた反応次元の曖昧さと選択された顔特徴の恣意性である。(1) 観察者による弁別と評定実験において怒りと幸福の程度という両極性の知覚的表現の次元が仮定されたが、このような値を両極とする単次元を仮定することには心理学的な根拠が乏しい。また、もしこのような次元を仮定しないならば、FLMP においてどのように反応が決定されるのかが説明されないことになる。(2) 実験において操作された特徴には Ekman & Friesen (1975) などを考慮に入れて多くの特徴のうち顔面の上部から眉と下部から口元の2つが選択されたが、この特徴選択は心理学的な根拠に基づくというよりも、FLMP を観察者による怒りと幸福の二肢選択、ないしそれらを両極とする評定に適用するための現実的な要請によるということが出来る。またこのような二肢強制選択事態では、観察者が実際に怒りや悲しみを知覚していたかどうかは明確ではなく、したがって選択・操作された特徴の処理と影響についての説明も多義性が残ると言わざるを得ない。

選択的特徴体制化モデルの提案

Hasegawa & Unuma (2008a) は、Ellison & Massaro (1997) における FLMP の妥当性の検証についての問題点をふまえて、単一感情の強さ、すなわち一方向の評価次元による知覚表情の測定をおこなった。またそこでは複数の選択肢への反応を前提とする FLMP は採用されず、単次元での観察者の反応値の予測値があらたなモデルによってもとめられた。具体的には LAM の前提から出発し、複数の顔特徴の線型一次結合で表情の知覚される強度が決定されることが仮定された。検討された顔特徴は、最近の行動的データ (Yamada, 1993; Shibui, Yamada, Sato, & Shigemasu, 2001) に基づき眉の傾きと目や口の開きぐあいという3つの特徴、およびそれらの間の空間的距離が取り上げられ、Huber & Lenz (1993) と同様の図式顔におけるこれらの特徴が操作された。一方、これらから知覚される感情の強さの次元としては怒りと悲しみが取り上げられ、その観察者による強度評定の実測値がモデルからの予測値と比較された。

ここで、操作された3つの各特徴は以下の2種類のコードによって数値化された。まず3特徴は絶対的大きさ(長さ)に依存しない値、すなわち眉毛の傾きはその角度の正接 (tangent) で、目と口の開きぐあいはそれぞれの横と縦の比で記述され、これらの値は相互に独立に変化した。このように数値化された変数によって知覚される感情強度を予測しようとするモデルは

特徴比率モデル (feature ratio model: FRM) とよばれた。これに対して、3つの特徴の絶対的な値、すなわち眉の傾きに対応する眉の上端から下端までの垂直距離、目と口それぞれの上端から下端までの距離 (高さ) を変数として取り入れたモデルは、特徴距離モデル (feature distance model: FDM) とよばれた。このモデルには、各特徴間の最短の空間的距離を変数とするモデルも含まれた。FRMは個々の特徴を個別に処理することを仮定する点で特徴分析モデル (feature analysis models) の前提に対応し、一方 FDM は特徴体制化モデル (feature configural models) に対応する。

重回帰分析によって観察者から得られた観測値にモデルからの予測値をあてはめた結果、知覚された怒りの強度については3変数を説明変数とする FDM が最も観測値に適合した。FRM および FDM の1変数および2変数モデルはいずれも3変数 FDM より低い適合度と評価された。さらに観察者全体のデータに関しては、3変数のうち眉毛の傾きに対応する変数が最も高い影響をもつことが示された。また観察者別にモデルをあてはめた場合にも、眉毛の傾きを変数に含む3変数および2変数の FDM が他のモデルよりも観測値に適合し、眉毛が最も重要な変数であった。これに対して悲しみの知覚された強度に関しては、怒りの場合と同様に3変数 FDM が他のモデルよりも観測値に適合したが、各変数への重みは怒りの場合と比較して眉毛の傾きの影響が低く、相対的に目の開きの影響が増加した。

これらの結果は、知覚される感情の強度が特徴間の関係を線型に加算するモデル (LAM) よりも、特徴に選択的な重みをつけて加算したモデル (weighted linear additive model: WLAM) によって説明されることを示している。また、この重みは知覚される感情によって変化すると言える。さらにそこで符合化される特徴は、特徴の個別処理 (FRM) よりも特徴間の空間的関係を含む情報の符号化 (FDM) をもとに処理されることが示唆される。また、特徴の線型加算モデルが知覚された表情の強度を説明したことから、表情の知覚がカテゴリカルな特性では説明できない連続性をもつことも示されている。

これらの知見は顔面の部分的特徴が表情の知覚を規定することを示すだけではなく、従来の特徴体制化モデルにおいて考慮されていない特徴選択的な体制化 (feature selective configuration) が顔特徴をもとにした表情の知覚の基礎にあることを示していると言えよう。すなわち、表情の知覚は単純な (選択的処理をふくまない) 特徴の体制化に還元できない特徴の統合過程に支えられており、その過程は特徴の選択的処理を含む選択的統合 (selective integration) の過程であると考えられることができる。Hasegawa & Unuma (2008a) はこのような統合過程をふくむ情報処理過程の全体像を、特徴の符号化 (encoding)、統合 (integration)、そして決定 (decision) の3段階として提案している。

さらに Hasegawa & Unuma (2008b) は、Hasegawa & Unuma (2008a) で用いられた図式的顔よりも現実の顔に近い材料を用いて WLAM の妥当性を検証した。テレビ放送番組から抽出された 100 の表情をもとに 28 の顔特徴が分析され、これらの特徴をもとに構成された WLAM が観察者によって知覚された表情の強度の実測値（評定値）にあてはめられた。その結果、怒りの強さの知覚では眉毛の動きと口の開きが重要な特徴であること、悲しみの強さにおいては口の開きおよび顔の下部の特徴が関わっていることが示された。この結果は、WLAM の妥当性を示すとともに、表情知覚の連続性を示唆するものと言える。

表情知覚のモデル化における今後の展望

本研究では、表情の知覚を説明しようとするモデルを比較した結果、顔面の視覚的特徴とそれらの空間的距離を変数とした線型加算が、個別の特徴の処理のみを仮定する特徴分析モデルや全体モデルよりも有効であることが示された。また、知覚される表情の強度に注目したことで、表情知覚の連続性が指摘された。今後、WLAM の可能性はさらに以下の二点において検討されることになるだろう。第一に、WLAM が仮定する特徴およびそれらの関係の情報への処理の加重、すなわち特徴の選択的統合の神経生理学的基盤が明らかにされる必要がある。第二に、本研究が取り上げた特徴からさらに生態学的に妥当な特徴へと WLAM を拡張する可能性を検討することである。

神経生理学的研究は、部分的特徴および特徴間の関係が表情および顔の同定において処理されるという仮説を支持する資料を報告してきた。すなわち、表情の知覚に対応する部位が上側頭溝（superior temporal sulcus: STS）において報告されてきた（Hasselmo, Rolls, & Baylis, 1989）が、一方でこの部位が目や口などの特定の顔特徴に選択的に反応することが明らかにされている（Perret, Rolls, & Caan, 1982; Perret, Mistlin, & Chitty, 1987）。さらに下側頭葉（inferotemporal cortex）のニューロンの一部が、顔特徴、たとえば目、口、眉などの間の距離とその組み合わせに応答することが報告されている（Yamane, Kaji, & Kawano, 1988）。これは顔の知覚的情報処理において特徴とそれらの空間的關係が処理されることを裏づけるものであり、特にいわゆる腹側皮質視覚路（ventral stream）において特徴間の関係が処理されていることを示唆している。

これらの知見に対して、本研究が取り上げたような特徴選択的な、すなわち個々の特徴の重みづけられた体制化に対応するような神経生理学的な基礎は未だ解明されていない。この点に関して、Moran & Desimone (1985) が外線条皮質（extrastriate cortex）、特に

V4あるいは下側頭葉における注意選択的な特徴の処理を指摘したことは示唆的である。さらに、図と地などの視野の体制化によって一次視覚野（V1）におけるニューロンの応答が変化することが指摘されており（文脈依存的修飾：contextual modulation）、この基礎には外線条皮質からV1へのフィードバックによる特徴処理の変化が仮定されている（Zipser, Lamme, & Schiller, 1996）。したがって、本研究が提案する選択的特徴体制化の神経生理学的基礎にはV1から外線条皮質までの双方向的な関与を考慮する必要があると考えられる。

一方、本研究が検討した顔面の視覚的特徴は静止した2次元画像上の特徴であった。この点では今後より現実的で生態学的な視覚的特徴と知覚される表情との関係が検討される必要がある。特に顔面の時間的な変化にともなう特徴の処理は残された重要な問題である。Edwards (1998) は表情変化の速さを操作して時間的情報の役割を検討した。また表情変化の系列を操作した Kamachi, Bruce, Mukaida, Gyoba, Yoshikawa, & Akamatsu (2001) は、顔特徴およびそれらの関係の処理の時間的特性を検討する方法を提案している。このような研究からさらに生態学的に妥当な視覚的特徴が抽出されモデル化されることで、表情知覚研究は新たな展開をむかえることになるとと思われる。本研究が提案した選択的特徴体制化のモデルは、そのような試みにおいても知覚情報処理の本質的な特性を表現するものとして発展することが期待される。

注

- 1) 本研究は、平成20年度川村学園女子大学教育研究奨励「物体知覚に関する計量心理学的研究と教育場面への応用（研究代表者 鶴沼秀行）」による補助を受けた。

引用文献

- Bruce, V. (1988). *Recognising Faces*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Brunswik, E., & Reiter, L. (1937). Eindruckscharaktere schematisierter Gesichter. *Zeitschrift fuer Psychologie*, 142, 67-134.
- Collishaw, S.M. and Hole, G.J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception*, 29, 893-910.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. John Murray, London.
- Duchenne de Boulogne, G.-B. (1990). *The mechanism of human facial expression*. Cambridge, England: Cambridge University Press. (Original work published 1862)
- Edwards, K. (1998). The face of time: temporal cues in facial expressions of emotion. *Psychological Science*, 9, 270-276.

- Ekman, P., & Friesen, W.V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Ekman, P., Friesen, W.V. & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings*. Elmsford, NY: Pergamon Press.
- Ekman, P., & Friesen, W.V. (1978). *Facial Action Coding System*. Consulting Psychologists Press, Palo Alto.
- Ellison, J.W., & Massaro, D.W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1), 213–226.
- Essa, I., & Pentland, A. (1994). A vision system for observing and extracting facial action parameters. *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR'94)*, Seattle, WA, USA, 76–83.
- Etcoff, N.L., & Magee, J.J. (1992). Categorical perception of facial expressions. *Cognition*, 44, 227–240.
- Hasegawa, H., & Unuma, H. (2008a). Roles of facial features for perceiving the intensity of facial expressions on schematic faces. Manuscript submitted for publication.
- Hasegawa, H., & Unuma, H. (2008b). Selective feature configuration and perceived intensity of facial expressions in realistic faces. Manuscript submitted for publication.
- Hasselmo, M.E., Rolls, E.T., & Baylis, C.G. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Experimental Brain Research*, 32, 203–218.
- Huber, L., & Lenz, R. (1993). A test of the linear feature model of polymorphous concept discrimination with pigeons. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology*, 46B, 1–18.
- Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S. & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, 30, 875–887.
- Leder, H., & Bruce, V. (1998) Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 51A(3), 449–473.
- Massaro, D.W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review*, 97, 225–252.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782–784
- Nosofsky, R. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance*, 17(1), 3–27.
- Perret, D.I., Mistlin, A.J., & Chitty, A.J. (1987). Visual neurones responsive to faces. *Trends in Neuroscience*, 10, 358–364.
- Perret, D.I., Rolls, E.T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47, 329–342.
- Reed, S. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382–407.
- Sayette, M.A., Cohn, J.F., Wertz, J.M., Perrott, M.A., & Dominic, J. (2001). A psychometric evaluation of the facial action coding system for assessing spontaneous expression. *Journal of Nonverbal Behavior*, 25, 167–186.
- Schwanger, A., Wallraven, C., Cunningham, D.W., & Chiller-Glaus, S.D. (2006). Processing of facial identity and expression: a psychophysical, physiological and computational perspective. *Progress in Brain Research*,

156, 321–343.

- Searcy, J.H., & Bartlett, J.C. (1996). Inversion and processing of component and spacial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4), 904–915.
- Shibui, S., Yamada, H., Sato, T., & Shigemasa, K. (2001). The relationship between the categorical perception of facial expressions and semantic distances. *Shinrigaku Kenkyu*, 72(3), 219–226.
- Tanaka, J.W., & Farah, M.J. (1991). Second-order relational properties and the inversion effect: testing a theory of face perception. *Perception & Psychophysics*, 50(4), 367–372.
- Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, 46A, 225–245.
- Yamada, H. (1993). Visual information for categorizing facial expression of emotions. *Applied Cognitive Psychology*, 7(3), 257–270.
- Yamane, S., Kaji, S., & Kawano, K. (1988). What facial features activate face neurons in the inferotemporal cortex of the monkey? *Experimental Brain Research*, 73, 209–214.
- Zipser, K., Lamme, V. A. F. & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, 15, 7376–7389.