

Abstract

Cognitive Processes of Emotion-Space Interaction in Facial Expression Recognition

Reiko SAWADA

Graduate School of Informatics

Department of Intelligence Science and Technology

Emotional facial expressions play a crucial role in social interactions. The process of recognizing them depends on the observer. When perceiving facial expressions, the observer not only perceives the face itself but also its location. Prior research has demonstrated that facial location influences facial perception and emotional recognition. Positional information, particularly in alignment with the positive-up association, affects facial emotion recognition. Therefore, spatial information is an essential factor in understanding facial emotion processing. This dissertation aims to elucidate how spatial factors—such as facial location or environmental spatial information—affect facial emotion processing.

The processing of emotional facial expressions is proposed to follow a sequential process: emotional facial features are extracted from visual input, analyzed to recognize the emotion, and utilized to infer the emotional state of the expresser. Based on this inference, the observer prepares and executes an appropriate behavioral response. While the spatial locations of facial expressions affect emotional processing, it remains unclear which stages of this process are influenced by spatial information.

The primary purpose of this dissertation is to determine how facial location affects the observer's emotional processing. To address this, we focus on two key challenges: (1) examining the influence of facial expression location on emotion recognition (Chapters 2–3) and (2) investigating the interplay between facial emotion and spatial information (Chapters 4–5). To explore these questions, we conducted psychological experiments that assessed the emotion-space interaction using behavioral measurements such as rating scores, accuracy, and reaction times. Additionally, we sought to clarify which stages of emotional processing are influenced by spatial information through the emotion-space association. To reveal the underlying cognitive processes, we applied the diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2008)—a standard sequential sampling model—to

reaction time data from correct and incorrect trials in binary choice judgments. In this model, evidence is continuously and constantly accumulated for making one of the two possible responses throughout the decisional process. We employed this analysis to decompose decision performance into three stages: (1) the pre-decisional process including attentional direction and encoding; (2) the evidence accumulation process for decision-making; and (3) the post-decisional process which involves response execution. These stages align well with the facial emotion processing model: the pre-decisional stage corresponds to the encoding of a facial expression, the decisional stage involves recognizing the expression and inferring the emotional state, and the post-decisional stage engages the execution of a behavioral response.

Chapter 2 investigated how facial location influences the behavioral performance of explicit and implicit emotional processing. Explicit emotional processing refers to tasks that require direct evaluation of facial emotions, such as emotional rating and classification tasks. In contrast, implicit emotional processing involves tasks unrelated to facial emotions, such as a discrepant face detection task. We found an emotion-space interaction in reaction times during the emotional classification task. Specifically, positive emotions were classified more quickly when facial expressions were presented in the upper spatial location compared to the lower one. However, no such interaction was observed in the discrepant face detection task. Instead, an upper spatial advantage was found for detecting target expressions regardless of their emotional valence. Additionally, the facial location did not influence the degree of perceived emotional positivity. These results indicate that facial location affects the speed of emotional classification in a manner consistent with the positive-up association.

Chapter 3 further examined the underlying cognitive process of emotional classification and discrepant face detection tasks using diffusion model analysis. This analysis revealed that the non-decisional process related to response execution was shorter for positive emotion classification when a happy face was presented in the upper spatial location compared to the lower one. In contrast, no evidence suggests that target location influenced the discrepant face detection process following the positive-up association. Thus, reaction times for explicitly recognizing positive emotions were faster when facial expressions were presented in the upper spatial location than in the lower one. This effect may be attributable to the facilitation of response execution, driven by the alignment between facial emotions and locations with the positive-up association.

However, when manipulating emotional and spatial information concurrently, it was unclear whether the upper spatial information influenced the recognition of positive emotions, or if positive emotions facilitated the recognition in the upper space. To address

this, Chapter 4 explored the interplay between facial emotions and spatial information by sequentially presenting emotional and spatial stimuli. We conducted two experiments using priming paradigms, in which the roles of emotional and spatial stimuli were reversed as primer and target. We examined both physical and conceptual aspects of spatial information. Alphabets presented in either the upper or lower locations of the screen were used as physical spatial stimuli, while shapes positioned at the center of the screen (identical to the position of the facial expressions) but conceptually representing upward or downward directions were used as conceptual spatial stimuli. The results indicated a bidirectional interplay between facial emotion and spatial information, aligning with the positive-up association. Positive facial emotions were recognized more quickly than negative ones when preceded by upper spatial primers, and upper spatial recognition was faster when preceded by positive emotional primers than negative ones. These reciprocal influences were observed for shapes representing vertical directions, but only when participants recognized spatial verticality spontaneously. This finding suggests that conceptual space, represented by symbols, is integrated into the connection between positive emotions and upper space. Moreover, the bidirectional influences were asymmetric, suggesting that different cognitive processes underlie each directional influence.

In Chapter 5, we estimated the cognitive processes underlying each directional influence using diffusion model analysis. The results showed that the priming effect of upper space on emotional classification was driven by the facilitation of evidence accumulation for recognizing facial emotions, consistent with the positive-up association. In contrast, the emotional priming effect on upper space recognition was attributed to an enhanced non-decisional process related to response execution, suggesting that emotional primers automatically activated spatial processing. These findings may reflect an asymmetrical relationship between positive emotions and upper space: spatial concepts associated with emotions are automatically processed during emotion recognition, whereas emotional associations linked to spatial concepts are not necessarily activated when processing spatial information.

The experimental findings consistently demonstrate that facial location and vertical spatial information from the environment influence facial emotion recognition, supporting a bidirectional yet asymmetric positive-up relationship. However, this influence was observed only when emotional or spatial information was explicitly processed. Diffusion model analysis further clarified how vertical spatial information affects the cognitive processes underlying emotional recognition. Specifically, an upper facial location enhances response execution for positive emotional faces, while prior

exposure to upper space facilitates the evidence accumulation process for judging emotional positivity.

Overall, these findings emphasize the importance of considering how observers perceive facial expressions—particularly, concerning facial spatial positioning and additional spatial cues within their visual environment. This observer-dependent perspective on emotional processing has significant implications for real-world communication both interpersonal interactions and human-robot interactions. For example, when a robot conveys friendliness through facial expressions, positioning its face higher than the user's viewpoint may elicit more efficient responses from the user. Additionally, robots and automatic emotion recognition systems could be designed to react more human-like by considering their eye level or directional information perceived just before a user comes into view.