The Author’s Gesture: The Camera as a Body in Wong Kar-wai’s *In the Mood for Love*

Jake Ivan Dole

In an early scene of Wong Kar-wai’s *In the Mood for Love* (2000), Li-zhen (Maggie Cheung) arrives at the apartment of neighbor Mo-wan (Tony Leung) to return a martial arts series. She is greeted by landlady Mrs. Suen and enters the room from the corridor to continue the conversation. Paradoxically, the camera that initially frames Li-zhen does not follow in her footsteps, but lingers in the corridor for an additional 14 seconds before a cut inside the room ensues. Originally positioned to Li-zhen’s right and by the side of the apartment’s door frame, the camera retreats subtly to the right, as if demonstrating an unwillingness to commit to Li-zhen’s goal (see clip below).

The camera’s lingering nature is surprising if one views the scene with a conventional frame of mind, because one will tend to expect the camera, as a formal component, to concern itself with the salient elements of the plot—Li-zhen and Mo-wan’s affair. Instead the motion of the camera and the decision not to cut demonstrate a counter-intentionality by not following Li-zhen’s enactment. The camera’s autonomy invites us to consider it a living, conscious being.
Unconventionally motivated camera behavior is sometimes interpreted as signifying the director’s artistic signature. André Bazin¹ and Alexandre Astruc² argued that film form is the language through which the artist expresses his thoughts and presence. When formal devices such as the camera are bared, and the narrational salience of classical storytelling is set aside, the director’s creative presence is presumably brought to the fore.

Wong Kar-wai’s camera work in *In the Mood for Love* has been discussed by scholars using a similar frame of reference, although more implicitly than explicitly. Nancy Blake attributes authorial emotions to the film’s use of a slow motion camera, writing that the stylistic decision reveals “a desire to immobilize a fleeing reality, to make time stand still (although) the camera can only ever render an image.”³ Blake adds that “Wong’s camera… lovingly dwells on Maggie Cheung, as she tries to come to an understanding of her husband’s desire.”⁴ Here Blake combines thematic context and elements of style to infer a particular set of emotions to the mobility of the camera, which becomes an avatar for a perceiving, feeling artist. In another reading, Rey Chow interprets the author’s pervasive gaze by reference to thematic content, style and socio-cultural context. She reveals that the setting of the film, Hong Kong in the 1960s, is a place of Wong’s childhood “remembered in oneiric images” that are “mediated by a particular consciousness.”⁵ Chow therefore attributes a nostalgia to the film’s mise-en-scène, a sensation projected by Wong’s mobile presence.

Wong’s camera also demonstrates a seeming corporeality, a physiology. In addition to its subtle resistance towards narrative goal-orientation, the camera lacks regimented motivation. It moves with subtle, and at times uncertain, intentions. The author we infer here is not only a social, formal or narrational construct, but furthermore a bodily one. The view of the camera as a corporeal presence was perhaps most finely articulated by Vivian Sobchack, who argued for an understanding of camera movement as an implicitly embodied subject—as an other who sees and expresses perception, and participates with the surrounding world.⁶ Sobchack writes that “camera movement echoes the essential motility of our own consciousness as it is embodied in the world and is able to accomplish and express the tasks and projects of living.” In other words, the camera is bodily in that it engages with the visual field, other bodies and objects in space, in a manner that compares with our own. It can perceive a range of tactile possibilities, direct itself towards objects and bodies, and make contact with objects and bodies.

Consistent with this thinking, whenever a camera is shown to project itself forward and towards an object, it enacts a familiar disposition and sensorimotor act—an intentionality that is satisfied through the movement of the body. This type of movement might communicate something narratively significant within a scene, but
could also generate empathetic attachment on the basis of bodily identification. Recent approaches in cognitive film studies have sought to couple the study of the sensing body and the sensorimotor experience (movement), with that of the cognitive activity of viewers.  

For example, in a case study of a scene from Alfred Hitchcock’s *Notorious* (1946), Vittorio Gallese and Michele Guerra argued that imagery of camera movement could induce in the viewer the imaginative sensation of motor mimicry within particular narrative contexts. When the film’s heroine Alicia (Ingrid Bergman) looks off-screen, the scene cuts to a tracking shot that approaches a desk on top of which lies a stack of keys. Earlier, with the aid of a flashback, it is established that Alicia intends to and succeeds in stealing the keys. The authors argue that when the forward tracking shot occurs in real time, Hitchcock’s camera simulates the forward motion of a body as it proceeds towards an object of the character’s desire, prepared to grasp it. Through a combination of physiology (camera’s forward movement), additional stylistics (such as cutting) and narrative context, viewers are invited to experience the movement and urge for tactile contact through Alicia’s point of view. Scholars in embodied cognition refer to this effect of motor mimicry as embodied simulation, which occurs when bodily enactments are observed by the viewer within her visual field and are resonantly felt through his or her still body. In the case of this scene, narrative context also nudges the viewer to associate the camera’s physiology with that of the heroine.

The missing element within this *Notorious* case study, however, is socio-cultural context. Gallese and Guerra analyze the sequence based on stylistics, narrative and physiology, and present an effective argument. However the notion of an author’s body represented by the avatar of the camera is a generally understood social construct that has roots in the ideas of Bazin and Astruc, as well as Dziga Vertov. To fully understand Wong’s camera as his body requires not only reference to its manner of movement (its physiology), or its narrative and stylistic function, but also additional reference to socio-cultural interpretive frames that the film projects and the viewer construes. And in order to fully appreciate the viewer’s conception of an author’s body (and projected emotions like “desire”), we need an approach that is epistemologically pluralistic.

In *In the Mood for Love*, Wong’s camera invites to be seen as a body and convinces as one for the following reasons: A) its physiology is visibly corporeal because it enacts particular bodily intentions and dispositions; B) it performs a narrative role that is consistent with pervasive cultural and theoretical assumptions about authors, directors and artists; C) the film’s setting and mise-en-scène function to prompt in certain viewers particular associations that can also be attributed to the camera’s movement; and D) certain viewers’ familiarity with Wong Kar-wai’s background will also affect their
understanding of the camera as the author’s body. Altogether, the more of these categories are satisfied in the viewing process, the more likely it is that the viewer will accept its role as that of an authorial body and presence. By representing himself via a camera, Wong invites the viewer to intersubjectively attend the film’s themes and imagery through his sensing and perceiving body.

The Camera’s Physiology

Anthropomorphic functions are frequently attributed to the camera in critical and academic writing, but with varying degrees of specificity. For example, David Bordwell commonly refers to steadicam and virtual motions of the camera in contemporary Hollywood as prowling, a verb that evokes the stealthy movements of a predator. However describing the camera in this way does not necessarily mean that one thinks of it as a body, nor that one is embodied by it. In order to establish an intersubjective relation between the viewer and the camera, there is a need for some clarity about the types of bodily cues (to borrow Gallese’s and Guerra’s term) that can be evoked by a camera and when.

Studies in embodied cognition and cognitive neuroscience about cinema inform that the viewer’s cognitive activity (e.g., story inference) is closely tied with the body’s interaction. In the process of watching a movie, the viewer’s body is activated in numerous ways, including the process of perceiving and reacting to bodies of on-screen characters. The perceptual encounter with on-screen bodies largely depends on a measure of “action understanding,” which ensues when the viewer observes intentional, goal-oriented actions. In other words, by seeing certain kinds of enactments performed by characters’ bodies, viewers emulate them in their active minds while the physical body remains still—what is normally termed as embodied simulation. Due to the presence of mirror neurons in the brain, the act of observing an action leads the observer’s brain to activate the same neural mechanisms that are normally trigged by performing the action oneself. Therefore a scene that shows a familiar intentional gesture, such a character squeezing a door handle, will trigger a response on the part of the viewer due to the inherent familiarity of the sensation.

Embodied simulations of this kind function without much strain or reflection, whereby we internalize the actions we see. Most of the time, we are not aware of the fact that we do so, meaning that the simulations we run in our minds are virtually invisible, or transparent, in their overlap with our conscious, reflective mental faculties. The feeling of the body therefore manifests itself nonconsciously, as second nature.
In order to think of a camera as a body, we need to consider whether its movements can effectively perform actions and dispositions that would elicit action understanding, and lead to embodied simulation. It is important then to demonstrate that the camera can enact goal-oriented movements and draw itself into intentional relations with surrounding space, in a manner that compares with a perceiving body. What then are the “bodily, tactile cues,” to use Gallese and Guerra’s words, that can be activated by a camera, and in this case by Wong’s camera specifically?^{17}

I suggest that we can look at camera movements in terms of particular action-based image schemas, which are basic abstractions from the body. Warren Buckland writes that schemata “represent the body in the mind, and make reasoning possible by providing a context for reasoning to operate.”^{18} Schemata then are mental structures that develop from fundamental bodily experience, and include examples like up-down, back-front, centre-periphery, part-whole, inside-outside, among others. Furthermore, basic action concepts such as grasp, put in, take out, crawl and hit get their meaning from our bodies and our ability to imagine enacting such tasks.^{19} The mirror neuron system enables every person to imagine acts like grasping and crawling, and is the very same capacity by means of which the viewer can simulate the actions of characters in the visual field. This prompts the question—how can the camera cue particular action structures in a convincingly bodily way?

We must admit that the camera can neither grasp, nor demonstrate apparent hand- or arm-based enactments in the conventional sense. One of the challenges in assigning the camera a sensorimotor role is its seeming lack of limbs and thus a limited capacity for transitive movement. Nevertheless, in the scene of Li-zhen’s visit, Wong’s camera communicates an action concept: it conceals itself slightly behind a wall. In fact, the camera’s tendency for concealment can be observed numerous times throughout the film, whereby it is seen positioning itself behind objects that partially block its view of the main line of story action. Viewers might be inclined to think of the camera as a lurking observer, perhaps shy or even voyeuristic, and periodically distracted. This type of inference results from encountering a familiar disposition of bodily concealment behind an object or barrier.

While certain objects conceal the camera’s view, others serve as temporary points of attention. Throughout the film, Wong’s camera is seen periodically readjusting, losing and regaining attention on objects in space. In doing so, the camera is almost never still, which communicates a seemingly drifting attention span. Even in his use of close-ups Wong slowly pans the camera, thereby subtly guiding the viewer’s perception towards its own movement. A couple of early close-ups in the film—that of a gift-wrapped box and a stack of books—are presented to the viewer for several seconds at a time and are ambiguous in terms of their narrative motivation, leading one to assume...
that the objects might be deeply, perhaps symbolically, meaningful to the author himself (See figures 1 and 2).

I propose that we can explain the close-ups from the embodied perspective: the objects have canonical affordances. Scholarship on embodied cognition has pointed out that, in certain instances, the mere act of looking at a manipulable object can anticipate potential action. Manipulable artifacts are those objects whose function anticipates guided action. For example, tools (e.g., a hammer) are meant to be grasped, a door handle is meant to be turned or squeezed, a button is meant to be pushed, etc. Wong’s close-ups, due to their persistent tendency for movement,
emulate the camera-body’s potential approach, and prompt the urge to reach out, touch the book and the box, and perhaps even unwrap the paper. In other words, the objects do not necessarily have to be understood in terms of their independent symbolic function, but rather as elements at hand—within the visual range and reach of a perceiving, mobile body. In stressing the mobile body’s perceiving role, Wong activates another familiar bodily disposition of proximity.

A somewhat similar disposition of proximity is evoked in Alfred Hitchcock’s *Marnie* (1964). In one climactic scene, the film’s eponymous heroine is seen stealing several wads of cash from the company safe and, in order to slip out unheard by the cleaning lady, she removes her shoes and quietly walks towards the exit. Hitchcock uses editing and close-ups to effectively draw the viewer’s range of knowledge onto one of the shoes as it gradually begins to slip from Marnie’s pocket. The intensity of the close-ups, combined with the disconnect between the viewer’s range of knowledge with that of the character, generates much of the scene’s suspense. However, a crucial component of the scene is also physiological: the close-up cues the viewer’s understanding of bodily proximity and the tactile familiarity with the manipulable artifact in the visual field. In other words, the close-up teases the viewer’s inclination to want to grasp the shoe and place it back into Marnie’s pocket (see clip below).

The crucial difference however between Hitchcock’s scene and Wong’s close-ups is that *In the Mood for Love* places visual emphasis on both the object and the perceiving, moving subject. The movements of Wong’s camera often communicate
intentions that are removed from the plot’s salient lines of action. This camera’s autonomy from classically motivated duty therefore draws the viewer’s eye to its own bodily mode of orientation in relation to the object. Although the Marnie scene will effectively trigger the viewer’s muscle memory in relation to the shoe, the camera in the scene, and throughout the film, is either static or committed to following the character.

**The Author’s Expressivity**

It is therefore possible to determine instances in which a camera can convey bodily intentionality, even in cases where its point of view does not belong to a character in the story. However attributing this body to the author’s point of view requires that the viewer make narrative and socio-cultural inferences as well. This includes narrative and genre expectations, general knowledge about authors and artists, as well as possible background knowledge about the director of the film.

Those who view *In the Mood for Love* as an art film will be inclined to see the camera’s counter-intentional movements as reflective of an authorial signature. The film’s key line of action, Li-zhen’s and Mo-wan’s affair, and certain activities of characters on-screen are periodically set aside in certain scenes. For example, during a brief scene in Mr. Ho’s office, the camera directs itself askew of the characters. At first we see Li-zhen speak on the phone, hang up and turn her body, but the camera ignores her movement and instead chooses to wander around the office, paying attention to various objects. Within the frame of expectations of classical cinema, Wong’s camera breaks with stylistic salience and demonstrates an alternative autonomy to that of the narrative. In such instances one is inclined to look at the author as the cause of the disturbance.

Per David Bordwell, works that subvert classical conventions of storytelling and style can be understood through the interpretive frame of “authorial expressivity,”23 whereby the auteur (usually the director) imposes on the work through his or her “overriding intelligence.”24 Such interpretations do not simply draw on the formal and physiological evidence, but are in fact culturally constructed. We tend to think of auteurs as somewhat manipulative, even arrogant, beings who wish to do more than simply tell a story. For example, we know that Jean-Luc Godard imposed his artistic signature on À bout de souffle (1960) by subverting classical continuity through jump cuts. Similarly, Alfred Hitchcock at times used a mobile camera to direct the viewer’s attention away from narrative salience, like in the case of *Frenzy* (1972) where his tracking camera abandons Robert Rusk and his victim by leaving the corridor of a building.
Thus whenever Wong’s camera subverts classical expectations by tracking or panning away from a line of action, the viewer will be inclined to attribute this behavior to the director, Wong. However unlike À bout de souffle, Wong’s intervention is both bodily and physiologically relatable—it enacts familiar dispositions. We can empathize with the author’s presence in the work through embodied simulation. Godard’s jump cuts do not offer the same kind of immediacy to the viewer’s experience. His body is absent and the resulting affect is minimal. Although we can read authorial expressivity to these formal devices, they do little to cue bodily engagement.

Conversely, Wong’s camera activates cognitive and physiological responses at the same time, requiring that the viewer relate back to cultural and narrative context in order to make sense of what the camera does. Cognitive scholars like Maarten Coëgnarts and Peter Kravanja, as well as Adriano D’Aloia and Ruggero Eugeni, stress that bridging the study of the body with that of cognition is necessary in order to accurately account for the viewer’s experience. Whenever the author’s presence is enacted through a body, and is absorbed and interpreted as such by the viewer’s body-brain system, what ensues is a mode of sensorimotor identification that has been typically haphazardly understood (for example, when ignoring socio-cultural and narrative context).

Furthermore, the ordinary understanding of authorship tends to evoke bodily metaphors. If I were to say that Wong manipulates or shapes his film’s imagery through the use of the camera, I would be referring to fairly routine terms of reference. However these terms also inscribe a bodily projection, by leading us to imagine the author’s activity through sensorimotor imagery, specifically action concepts like shaping, manipulating and molding. Cognitive linguist Mark Turner made this type of observation when writing about the use of sensorimotor imagery in the writings of Marcel Proust, whose prose often relates the body to the narrator’s experiences.

In the opening volume of À la recherche du temps perdu (1913), Proust recalls his youth by conceptualizing his memories through action concepts. He writes, “My mind, striving for hours on end to break away from its moorings, to stretch upwards so as to take on the exact shape of the room and to reach to the topmost height of its gigantic funnel, had endured many a painful night.” Proust, the author and narrator, recalls his dream by physically shaping it with his body, stretching his joints to extend his body up and then reaching, perhaps with an extended arm. The imagery is extremely vivid due to the specificity of the body’s intentional movements and the detail of the prose.

Turner terms such bodily projections through the concept of the thinker as a mover and manipulator (otherwise termed THINKER IS A MOVER AND A MANIPULATOR in all caps). A fictional work about a narrator’s journey of the soul is often understood in
terms of such a projection, whereby the experience of conjuring a memory or constructing an imagined space is made to be figuratively corporeal. Unlike literature, movies create visual conditions for embodied simulation, but like literature they also rely on the viewer's understanding of many such unspoken cultural conventions that are stored in the recesses of our minds. Thus, in addition to being known as sources of an overriding intelligence, authors (like thinkers and like storytellers) tend to impose themselves on their work physically and, in doing so, place their readers and viewers into frames of embodied identification.

The resulting effect is different than that of a point of view shot or other kinds of identification with characters. The author’s grasp transcends the diegetic limits of the story, and demonstrates a seemingly unlimited power to manipulate the work’s formal contours. Therefore, to identify with Wong’s camera as the author’s body means to engage with the film in a metafictional way, whereby the film becomes the author’s raw materials at hand, while the viewer is made to feel the manipulation of that material with his or her body.

**The Body in the Setting**

Existing scholarly interpretations of *In the Mood for Love* tend to see Wong Kar-wai’s presence in the film as that of a nostalgic, somewhat Proustian, dreamer who meditates on his past. The distracted camera and its fascination with objects is interpreted by some writers in relation to the film’s setting, a 1960s Hong Kong of the director’s childhood.

For example, Nancy Blake refers to the film’s camera as voyeuristic, due to the aforementioned tendency for concealment as well as Wong’s background. Rey Chow uses a similar frame of reference when she describes Wong’s attention to the mise-en-scène as dream-like and nostalgic. Blake and Pam Cook write that viewers who share Wong’s cultural and generational background are likely to remember 1960s Hong Kong as a country that underwent a transformative transition towards capitalism. Wong’s desire to recreate a non-recoupable time and place of his childhood memories is also documented by Cook, and thus explains his decision to draw the viewer’s attention to memorabilia and period styles. The ordinary household items that are emphasized in the film through the stylistic decisions of panning the camera or cutting to a moving close-up, take on a meaningful resonance in light of this cultural context. Cook adds that the film is “littered with memories of personal significance,” producing “a kind of meditation on the passage of time.”
Therefore to viewers who share or know about Wong’s background, a bodily identification with the camera resonates more greatly due to this shared history and memory. The camera’s disposition towards objects, manipulable or otherwise, begins to evoke particular emotions. The invisible, distracted witness thereby invites that viewer to share a journey and relive familiar objects, colors and textures through his or her body. Wong thus places his avatar camera into a state of intersubjective relations with the viewer, and even more deeply with viewers who share this historical memory.

Chow writes that “for audiences already acquainted with the Hong Kong of the 1960s, these ethnographic details arguably constitute a kind of already-read text, one that evokes… the sense of community that has been but no longer is.”35 Blake echoes this insight when she writes that the camera “reflects a desire to immobilize a fleeing reality, to make time stand still, an effort to understand.”36 The word “desire” here is not simply used in a psychoanalytic way, but in an effort to articulate the same basic state of emotions discussed by Chow. Both scholars reflect on the respective roles of the camera and the mise-en-scène in relation to cultural memory, but only loosely make the connection between cognition and physiology. To me, this connection appears to be quite crucial, because the empathetic experience constructed by Wong here is based on the congruence of styles and the reactions they are meant to evoke. The film’s meditative experience is first made possible due to the bodily presence of a meditator (the camera) who actively and intentionally explores the film’s spaces with his body. The viewer’s cognitive inferences (about authors and cultures) fill in the rest of the information and appropriate emotional meaning (such as nostalgia or desire).

Therefore, Wong’s camera can be grasped within a pluralistic understanding of cinematic empathy, whereby bodily physiology, narrative context and socio-cultural knowledge (of the setting and the author) overlap to produce a desired emotional response.

**Conclusion**

To quote D’Aloia and Eugeni, embodied cognition stresses that the viewer’s cognitive activity depends on those experiences “that come from having a body with various sensorimotor capacities (that) are themselves embedded in a more encompassing biological, psychological and cultural context.”37 In other words, in order to understand cinematic empathy, we need to consider the overlap between the viewer’s physiology and cognitive activity in relation to both narrative and socio-cultural contexts. Up to this point, it has been fairly common to discuss the camera as a body while concentrating on physiology (e.g., Vivian Sobchack’s work). Conversely, traditional cognitive scholarship sometimes sidesteps issues of the body in favor of activity of pure
narrative inference (as documented by D’Aloia and Eugeni). Within both approaches, it is not uncommon for socio-cultural context to be left out entirely. However this can be remedied with a more pluralistic approach that sees these different components as mutually inclusive.

With these various factors in mind, we can look at camera work in a film like In the Mood for Love as a convincing representation of an author’s body, whose perspective the viewer is positioned to simulate. Although it is normal to think of the author as a cultural construct, particularly in light of theories that push against notions of an auteur, I believe that particular formal inscriptions (such as ways of moving the camera) can create absorbing conditions for intersubjective embodiment. In Wong’s film, the camera’s bodily actions cue the viewer’s physiological response, which prompts cognitive inferences based on narrative comprehension and socio-cultural understanding. Even though one cannot isolate the author strictly within the camera (as a physical thing or as a visual manifestation), its movement inscribes a body and creates the conditions within which the viewer can flesh out a relationship with the meditating author. Within this understanding, Wong’s camera lays a foundation within which an experience of intersubjective attending, emoting and feeling is made possible.

Jake Ivan Dole is a PhD candidate in Moving Image Studies at Georgia State University, where he researches embodiment with visual media by combining phenomenological, cognitive and historical approaches.

Notes


4 Ibid., 349.


9 Ibid., 201.

10 Soviet filmmaker Dziga Vertov theorized about camera work through the concept of the *kino-eye*, a mechanical eye, as an extension of the body which, at the same time, could surpass bodily limitations. In his writings, he sometimes used anthropomorphic terms to describe the *kino-eye*’s movements, such as when he writes “I draw near, then away from objects, I crawl under, I climb onto them.” See Vertov, Dziga. *Kino-eye: the writings of Dziga Vertov* (Univ of California Press, 1984), 17.


14 Gallese and Guerra, “The Feeling of Motion,” 108.


17 Gallese and Guerra, “The Feeling of Motion,” 106.


Ibid., 59.


Turner, *The Literary Mind*, 44.

Blake, “‘We Won’t Be Like Them,’” 352.

Chow, “Sentimental Returns,” 646.

Blake, “‘We Won’t Be Like Them,’” 341.


Ibid., 4.

36 Blake, “‘We Won’t Be Like Them,’” 352.