The Philosophy of Computer Games Conference, Madrid 2012

Through the Image: Framing Videogame Play

Geert Bruinsma

Introduction

The importance of the concept of action in the analysis of videogame play is widely recognised and has emerged in many different formulations, ranging from interactivity to ergodic action to ludic action (Aarseth 1997; 2001; Galloway 2006; Mukherjee 2008). However, in the process of making a claim for the videogame's interactive nature, it has been argued that the experience of videogame play is not primarily visual but kinaesthetic, functional and cognitive (Aarseth 2004: 52). The corollary of this argument is that to emphasise the visual aspects of videogame play would be short-sighted as it would run the risk of obscuring the interactive nature of playing a videogame.

This paper aims to respond to these theses by introducing a methodological framework for the analysis of the nature of action in videogame play which takes the videogame image as its point of entry. Drawing on the phenomenological approach to action and perception by Maurice Merleau-Ponty and Alva Noë and Gilles Deleuze's concept of the movement-image, this framework will both be sensitive to the embodied nature of videogame play and offer possibilities for a nuanced understanding of action in videogame play. The paper will not only set forth an analytical model but it will also put this model to the test by analysing *Doom 3*, *Fallout 3*, and *Shadow of the Colossus*.

To see is to act

As the etymological root of "video" suggests (from the Latin videre: to see), visual perception functions as the prerequisite of videogame play; turn off the screen during a given gaming session and the effects are devastating to your game. This is not to say that the kinaesthetic, functional and cognitive facets of videogame play are of lesser relevance, but the physical and mental activity of pushing knobs on a controller only becomes meaningful in relation to the movements of the videogame image. In order to reveal the relevance of the screen and its images to the experience of videogame play, we must conceive of the player as an embodied subject involved in a physical and cognitive activity in which visual perception is key. Moreover, from a phenomenological point of view, physical action in videogame play – i.e. the pushing of knobs – and visual perception become indistinguishable. Perception is a form of physical action which is encompassed precisely by the kinaesthetic dimensions of videogame play.

> "According to the phenomenology of Merleau-Ponty, it is precisely through the body that we have access to the world. Action and perception are intertwined. In this notion, the concept of "flesh" becomes relevant. Merleau-Ponty uses the

> word ‘flesh’, as the domain in which experiences exist. Experiences are the mode of functioning by which we, inevitably, participate in the flesh. In terms of “the flesh” we are able to have direct, immediate contact with others and the world. My body is not able to forget its flesh. Although not always consciously, my body is always present and is involved in every action I undertake.” (Hermans 2004; 2)

Still, playing a videogame puts different physical demands on the player's body than, say, an actual game of soccer does. Playing a game of *Pro Evolution Soccer* 6 on my PC does not have me run, sprint, tackle or kick a ball; it has me sitting in front of the screen tapping away on the keyboard. But the concentration of the player is not geared towards the pushing of keys on a keyboard *per se*; the attention of the player is directed primarily to the screen.

It seems trivial to point to the importance of visual perception to the concept of action; given the integrated state of human senses, the visual tends permeate most of our every-day interactions with the world. From opening and walking through a door to pouring a glass of milk. Moments of audio only or haptic only are quite rare. In this sense, most of our daily actions do not exclude the visual. The danger, however, of compartmentalising the different facets of videogame play into discrete elements such as the visual, the kinaesthetic, the functional and the cognitive resides in that it tends to reduce complex physiological and biological processes to objective “functions”. Perception in this sense, becomes passive and immediate, much like the computer's blinking cursor waiting on you to type a command on the keyboard to which it instantaneously responds. But as theorist of visual perception James Gibson aptly points out: “The eye is not a camera that forms and delivers an image, nor is the retina simply a keyboard that can be struck by fingers of light” (Gibson 1979: 61)[1]. According to philosopher Alva Noë, visual perception does not merely *permeate our interactions* as an image being shoved under your nose; “perception is not something that happens to us. It is something we do” (Noë 2004: 1).

Noë, following the work of Merleau-Ponty, criticizes this so-called snap-shot conception of visual experience, by which seeing the world is like having detailed pictures of the world in mind. Discussing the work of Descartes on the biological and mechanical workings of the eye, Noë explains how the snap-shot conception holds the idea “that vision starts with retinal pictures that are transformed into better internal pictures that give rise to experiences with picture-like content” (ibid.: 39). However, in a phenomenological approach to visual perception, perception becomes more than the cognitive creation of an internal image based on a less perfect retinal picture. Instead, we perceive of the cat sitting behind a picket fence as indeed a whole cat and not a disjointed set of strips of cat because “[m]y sense of presence of the whole cat behind the fence consists precisely in my knowledge, my implicit understanding, that by a movement of the eye or the head or the body, I can bring bits of the cat into view that are now hidden” (ibid.: 63-4). To look at something takes up the entire body and our knowledge thereof; it is dependent on motor and muscle activity and as such it is

1 “The organism cannot properly be compared to a keyboard on which the external stimuli would play and which in their proper form would be delineated for the simple reason that the organism contributes to the constitution of that form. When my hand follows each effort of a struggling animal while holding an instrument for capturing it, it is clear that each of my movements respond to an external stimulation; but it it also clear that these stimulations could not be received without the movements by which I expose my receptors to their influence” (Merleau-Ponty 1965: 13).

kinaesthetic in nature. Moreover, visual perception does not deliver us a mere picture-like experience of the world, instead "we take ourselves to be situated in an environment, to have access to the environmental detail as needed by turns of the eyes and head and by repositioning the body" (ibid.: 59).

In this regard, an understanding of perception as a mode of action, as a physical *and* cognitive activity becomes crucial to a theory of action in videogame play. To put it crudely: while playing a videogame, to look is to take action. Visual perception functions not only as a prerequisite to the player's ability to play the game, but as a constitutive constant throughout videogame play. And even though playing a videogame in a sense immobilizes its player – keeping her sat in a chair, fingers on the buttons, and eyes glued to the screen – the ability to take any action relevant to the process of the videogame depends without interruption on an embodied activity involving the visual perception of the image on the screen.

The videogame and the image

A moment of caution is however warranted when addressing the videogame image; the videogame image is the visual expression of a computer program and as such it differs from more traditional images such as photo's or films. And if the videogame does provide us with moving images, then these images are not just there to be gazed and pondered upon; there are there to be acted upon. The images of the videogame *Tomb Raider*, for instance, do not present us the figure of a athletic young woman in the same way that the feature film does. In contrast to the inaccessible Lara on the film screen, the Lara on the videogame screen stands under direct control of the player. And by that the videogame image showing Lara takes on the function of a control panel. As researcher of new media Lev Manovich explains:

> "[I]nterfaces try to balance the concept of a surface in painting, photography, cinema, and the printed page as something to be looked at, glanced at, read, but always from some distance, without interfering with it, with the concept of the surface in a computer interface as a virtual control panel, similar to the control panel on a car, plane, or any other complex machine" (Manovich 2001: 91-2).

The videogame image as an interface can thus be said to oscillate between a visual representation and an interface-image. And the videogame image then operates between the code and the player by offering at once a visual representation *and* the means to control this representation. Instead of staying on the level of visual representation, we expect to go "into" the image by utilising it as a interface-image (ibid.: 290); the interface-image allows the player to take action over, or within the image. It is through the function of the interface-image that the player and the videogame image are integrated in the cybernetic feedback loop that constitutes videogame play. It is this cybernetic loop which again highlights the physical bond between the player and the videogame.

In some videogames, the control panel-dimension of the videogame image is made explicit. The game *Splinter Cell: Pandora Tomorrow*, for instance, has the player engage with in-game computer's, using the conventions of the Graphical User Interface to find information and read through e-mail conversations looking for clues (fig.1a). The interface-image is however made most explicit in the keypad door locks featured in *Splinter Cell* (fig.1b); the videogame

a)

b)

c)

d)

Figure 1. Screen shots from *Splinter Cell: Pandora Tomorrow* and *Hitman: Blood Money*.

image here takes on the shape and workings of an actual control panel. But also in less clear-cut cases, the interface-image comes to the fore wherever the visual appearance is coupled to a procedurally constructed mode of engagement, such as the visual appearance of an opponent in *Hitman: Blood Money* and the procedural representation that allows you to shoot him (fig.1c). In this sense, the cross-hair and the opponent function in much the same way as the buttons on the keypad and the pointer used to press them; they are represented both visually and procedurally.

The control panel-dimension of the videogame image is one end of the scale, the visual representation is the other. In videogames, the visual representational dimension is often dominant in those instances in which the player is not afforded any gameplay activity, such as pre-rendered cut-scenes, loading screens, or stat screens. These types of images differ from those mentioned above in that they go without the procedural representation that grands them real-time responsive behaviour. Figure 1d, however, shows an instance in which the visual representation stands out from the interface-image during player controlled gameplay; on the left side you see a ravine, with a ledge leading down to the bottom of it. Walking down this ledge is however not as exciting as it may seem, because you simply cannot step over the

edge and fall into the ravine. It is as if there is an invisible wall alongside the ledge with a picture of a ravine painted on it; it looks like a ravine, but it does not behave like one. Thus the ravine is represented visually but not procedurally.

Deleuze's movement-image

The videogames discussed in this paper are first- and third-person videogames, and in this respect de connection to Deleuze's cinema theory becomes visible. These types of videogames employ three-dimensional graphics and a dynamic frame to emulate what Michael Nitsche describes as a "virtual camera" (Nitsche 2008: 77). The correlations between the camera and the virtual camera centre on the aspect of framing; both types of camera involve the dynamic framing of movements and actions. Moreover, "the virtual camera is crucial in opening up the possibility of three-dimensional space, and allowing a character to move through those spaces" (Tavinor 2009: 66).

With regards to the framing function of the film camera, Deleuze identifies the frame as "a relatively closed system which includes everything which is present in the image – sets, characters and props" (Deleuze 1986: 12). This frame is a relatively closed system because the frame always operates in relation to what Deleuze calls the out-of-field; "[t]he out-of-field refers to what is neither seen nor understood, but is nevertheless perfectly present" (ibid.: 16). In the cinema image the out-of-field is made explicit and dynamic, as the moving camera and the possible reframings enabled by montage can bring into view that which was not in the image earlier. As Deleuze explains:

> "[T]he out-of-field already has two qualitatively different aspects: a relative aspect by means of which a closed system refers in space to a set which is not seen, and which in turn can be seen, even if it gives rise to a new unseen set, on to infinity; and an absolute aspect by which the closed system opens on to a duration which is immanent to the whole universe ..." (ibid.: 17).

The relative aspect of the out-of-field is the aspect which creates a concrete space through the relations between one frame and the next, be it through montage or by the moving camera. The absolute aspect of the out-of-field is most visibly achieved through images of disrupted space, joined images of distinctly different spaces with little or no relative spatial relation among them. This would suggest that the experience of cinematic space created through the relative out-of-view depends upon our bodily experience in, and of actual space.

It is the distinction between the relative out-of-field and the absolute out-of-field which leads Deleuze into a discussion of the shot, for it is in the shot that the movement of the mobile camera appears. The shot can be defined "as the intermediary between the framing of the set and the montage of the whole" (ibid.: 19). Hence, Deleuze maintains that "[t]he shot is the movement-image … it relates movement to a whole which changes, it is the mobile section of duration" (ibid.: 22).

The shot, or movement-image is however characterized by a number of distinct images, or avatars as Deleuze describes them, namely: the perception-image, the action-image, and the

affection-image[2]. And even though these avatars are described as images, and are even exemplified by Deleuze by distinct types of framing (long shot, medium shot, close-up), it is paramount to conceive of them as moments integral to all movement-images. Moreover, the breaking down of the cinema image into these three avatars by no means implies a temporal order among them. To Deleuze images are in and of themselves flowing matter, in constant flux or movement and to cut up and quantify movement would be to nullify movement all together. The identification of a perception-, affection-, and action-image is made for analytical reasons only.

The perception-image is the moment of seeing. It is the initial moment and remains present all during the duration of the movement-image. And while the continuity of perception seems superfluous to stress "if the world is incurved around the perceptive centre, this is already from the point of view of action, from which perception is inseparable" (Deleuze 1986: 64). Perception in this sense is not an abstract and passive registering of stimuli but always a combination of sensory and motor activity. To look at something is already an action. One can for instance choose to look away or close one's eyes. In the case of the videogame image it becomes most apparent how perception already entails action; to see the alien during a game of *Duke Nukem 3D*, entails the act of shooting it or being shot by it yourself.

The action-image, the second type of movement-image, follows naturally from the perception-image. According to Deleuze, the action-image is no longer concerned with elimination, selection or framing, but by the "incurving of the universe": "By incurving, perceived things tender their unstable facets towards me, as the same time as my delayed reaction, which has become action, learns to use them". (ibid.) In contrast to the perception-image, which is exemplified by Deleuze by the long-shot, the action-image is the most dominant type of movement-image in the medium and medium-long shot. This difference already points to the incurving, as the medium shot fills the frame with the action and rids it of the excess eye candy of the long-shot; here perception is subordinated by action.

Between the initial moment of perception and the "delayed reaction" of the action-image comes the affection-image. Deleuze describes the affection-image as "that which occupies the gap between an action and a reaction, that which absorbs an external action and reacts on the inside" (ibid.: 217). It is the moment where perception turns to focus on the action and where the reaction is delayed. It is, in the context of for instance the tennis videogame, the moment when you are fully focussed on the ball soaring towards you, in a split second contemplate whether to use your forehand or backhand. Deleuze illustrates the affection-image through two types of examples from the cinema: the close-up and the "any-space-whatever," Deleuze's concept of undetermined and fragmented space.

In the close-up (typically a close-up shot of the face, thought this need not necessarily be the case), a manner of framing only a portion of the whole body in such a way that the face appears relatively immobile, "[t]he moving body has lost its movement of extension, and movement has become movement of expression. It is this combination of a reflecting, immobile unity and of intensive expressive movements which constitute affect" (ibid.: 87). In other words: compared to the medium shot of the action-image, which frames for instance a

2 The avatars of the movement-image are, however, not limited to these three images, as Deleuze also distinguishes the impulse-image (which acts as an intermediate between affection and action), the reflection-image (which operates between action and relation) and the relation-image.

fistfight between two characters in a way that makes it unmistakable to perceive the action – an arm extents itself and the fist of one character comes into contact with the chin of the second one – the close-up presents only micro-movement – the raising of an eyebrow for example. Where the punch in the fistfight is unambiguous and immediately perceived and understood, the raised eyebrow raises questions as to what it might mean. The micro-movement of the raising of an eyebrow are then not movements of extension, but of expression, meant to designate a possible internal state of the character in question which is not directly perceivable.

If the close-up generates affect by de-contextualising the face, the "any-space-whatever" operates as to de-contextualise space itself:

> "Any-space-whatever is not an abstract universal, in all times, in all spaces. It is a perfectly singular space, which has merely lost its homogeneity, that is, the principle of its metric relations or the connection to its own parts, so that linkages can be made in an infinite number of ways. It is a space of virtual conjunction, grasped as a pure locus of the possible" (ibid.: 109).

As every movement-image will contain a perception-image and an action-image, so does the affective enters the movement-image through the any-space-whatever. For the edges of the image framing a certain portion of space always imply an as-of-yet-unforeseen relation to a subsequent framing of a portion of space. It is this unfulfilled promise of a possible linkage of one frame to the next which renders the movement-image its affective dimension. In the any-space-whatever, "space becomes tactile, as if the eye were a hand grazing one surface after another without any sense of the overall configuration or mutual relation of those surfaces. It is a virtual space, whose fragmented components may be assembled in multiple combinations, a space of yet-to-be-actualized possibilities" (Bogue 2003: 80).

From cinema to videogame

But videogames are no cinema. And a major problem arises when Deleuze collapses the significance of cinematic movement and images back onto the very images themselves. His cinema theory is a theory of the *cinematic image*, not that of the *cinematic experience*, and perception, action, and affection are construed as properties of the images themselves. So how is a theory of the cinematic image useful in understanding the experience of videogame play? The movement-image can help us understand the nature of the all pervasive action in videogame play, it helps to put it in to perspective and adds dimension to the concept of action. Even though the experience of videogame play is dependent on, if not constituted by the player's actions, videogame play can never merely be action. But, in order for the concept of the movement-image to migrate from the cinema to the analysis of videogame play, the properties of perception, action, and affection have to be opened up, and examined in their explicit relation to the embodied experience of playing a videogame.

As said, the connection between Deleuze's movement-image and the first- and third-person videogame becomes most apparent at the level of framing, yet here too, interesting differences appear. According to Deleuze, "[a]ll framing determines an out-of-field … There is always a thread to link the glass of sugared water to the solar system, and any set whatever to a larger

set" (Deleuze 1986: 16). The videogame, generally speaking, has a lot less potential to make abstract combinations between one framed set and the next. It is the player's actions which put the frame into motion and make the connection between one framed portion of space and the next. Thus, the framing function of the image of the 3d videogame is prone to the relative out-of-field, and the image in 3D videogame play operates, to a large extend, as a shot rather than a edited sequence. The dynamic frame of the virtual camera mostly establishes relative and measurable relations among the different framings, opening up on a homogeneous and actual space. Compared to the film director, the game designer then can not in the same degree permit the space to be dislocated from framing to framing. As theorist of visual culture Mike Jones explains:

> "The process of designing and producing a 3D game is both aesthetically and practically that of creating a macro-mise en scene containing the entire imaginary world. During game play individual frames will be 'composed' by the camera/player [through the workings of the virtual camera] but the larger macro-mise en scene remains fully intact and the player/viewer's awareness of it as a composition is never diminished" (Jones 2005: 1).

In the framing of the 3D videogame the affective already enters the experience of videogame play through the experience of space. Analogue to Deleuze's description of the close-up, the macro-mise en scene becomes de-contextualised through the micro-movements of the frame. The player only gets to see a portion of the macro-mise en scene through the frame at each given moment, having to relate each framing to a conceptual understanding of the macro-mise en scene. In this sense, the spatial understanding of a given frame stays with the player even after it has been replaced by the next framed portion of space. "With a first-person perspective, the spatiality of the world becomes far more *experiential* for the player: there is a 'near' and a 'far', an 'above' and 'below', and a 'behind'. Exploring virtual space is increasingly an experiential negotiation ripe with affective potentials" (Shaw and Warf 2009: 1339).

We will return to affect later on, but for now let us examine perception in Deleuze's movement-image and in videogame play. Paolo Marrati explains how for Deleuze,

> "[p]henomenologically oriented approaches cannot account for what belongs to cinema itself insofar as they retain subjective or "natural" perception as the model of reference, whereas the specificity of cinematographic perception lies precisely in the fact that it cannot be referred back to any subjective center" (Marrati 2003: 2).

The perception of the 3D videogame image, however, is characterised precisely by that; the ability to be referred back to a subjective centre. In classical narrative cinema, such subjective shots are scarcely employed (Galloway 2006: 68), and the cinematic perception – that which the camera brings into view – and the subjective perception of the spectator appear as two separate and disjointed visions. What the 3D videogame image shows, by contrast, is highly subjective; it monopolises subjective framings to display at once what the player-character 'sees' as well as what the player sees.

In his approach to the movement-image, Deleuze favours the cinematic perception, as it is the mobile camera which allows the cinema to present "the pure vision of a non-human eye, of an

eye which would be in things" (Deleuze 1986: 81). This is not to say that Deleuze finds cinematic perception to be totally devoid of access to subjective perception: "[I]s not the cinema's perpetual destiny to make us move from one of its poles to the other, that is, from an objective perception to a subjective perception, and vice versa?" (Deleuze 1983: 72). Yet it is Deleuze's notion of subjective perception in phenomenology which proves somewhat problematic: "What phenomenology sets up as a norm is 'natural perception' and its conditions. Now, these conditions are existential co-ordinates which define an 'anchoring' of the perceiving subject in the world, a being in the world, an opening to the world" (ibid.: 57).

The anchoring of the subject through natural perception in this understanding, becomes much like the way in which the three-dimensional painting or the photograph captures its viewer. The central perspective of these types of images organize the perceptive field of the viewer in such a way that the experience of perspective and space only opens up once the viewer is immobilized, frozen in a singular position. As an anchoring device for the subject in the world, natural perception takes the particularized and analytical perception of the fixed gaze as its model. In the phenomenological approach to perception, however, such detailed vision

> "appears when, instead of yielding up the whole of my gaze to the world, I turn towards the gaze itself, and when I ask myself what precisely it is that I see; it does not occur in the natural transactions between my sight and the world, it is the reply to a certain kind of questioning on the part of my gaze, the outcome of a second order or critical vision which tries to know itself in its own particularity, of an 'attention to the pure visual'" (Merleau-Ponty 2002: 263).

Merleau-Ponty rejects the conflation between the analytical and immobilized perception of the fixed gaze and the more holistic notion of natural perception. Rather than serving as a model for natural perception, the fixed gaze can be understood as an active attitude towards this very embodied state of unexplicated yet continuous perceptual activity.

When turning to the embodied notion of videogame play, this distinction between the continuously active, embodied model of natural perception of Merleau-Ponty and Noë and Deleuze's more rigid conception of natural perception, becomes interesting. The videogame image, by virtue of the framing determined by the physical dimensions of the screen and the framing workings of the virtual camera, already presents a circumscribed and (relatively) fixed gaze for the player to control. We might assume that a critical gaze would dominate the typically highly concentrated perception of the videogame player during play; but this alone can not render the rigid, Deleuzian conception a suitable model for perception in videogame play. For it seems unlikely to conceive that a player can meet the unrelenting fixation of the videogame image's framing with an equally unfluttering gaze demanding a constant amount of concentration from the player[3]. On the contrary, perception in videogame play is governed by a Merleau-Pontian notion of natural perception, to which the frequently employed fixed gaze appears as the outcome of a second order or critical vision.

As is the case with the perceptual activity of the player towards the videogame image – the

3 Natural perception according to Merleau-Ponty, is relatively stable, and is achieved "with our whole body all at once, and which opens up on a world of inter-acting senses;" whereas the fixed gaze, or discrete "sensory experience is unstable," depending on a distinct and forced sensorimotor effort of isolated sensory activity (Merleau-Ponty 2002: 262).

perception-image, in Deleuze's discussion of the movement-image, flows naturally into the action-image. Still, the videogame image and the cinema image operate in different ways. In the videogame image, perception extends itself not just to a depiction of action contained in the image, but into the active and constitutive bodily involvement of the player with the displayed action. As such, not only is the perception-image in the videogame highly subjective, so are the actions contained in the action-image. You – the player – push the button and as such experience the visual representation of the action, or movement in the videogame image as an extension of your pushing of the button. The videogame image is characterised by what cultural theorist Andrew Darley calls vicarious agency, "the impression of being enabled to act within and upon the world one gazes upon" (Darley 2000: 160). With the 3D videogame,

> "[t]he semblance of realistic spatial orientation is maintained from the cinema aesthetic, but at the same time it is heightened both by the capacity of the computer to model three-dimensional space and by the control one is given to determine where one goes and what one does" (ibid.: 159).

In this regard, the action-image takes on a different function in the context of the videogame image. In the videogame image, the action-image will always be connected to the player's actions through the interface-image. In other words: in the videogame image, the action-image operates between the interface-image and the instantly, yet delayed flash of representation of movement.

What Deleuze has done with his reduction of the cinema to the cinematic image – i.e. the construction of perception, action, and affect as properties to this image – is to subsume action and affection as subcomponents of the visually perceivable images, or as subcomponents to perception. Yet, Bergson has treated affection as an independent modality in its own right differing from perception. According to Bergson, the image that is my body, "I know from within, by sensations which I term affective, instead of knowing only, as in the case of the other images, its outer skin" (Bergson 2002: 114). In effect, then, "Deleuze has reduced affection to a formal process of technical framing, and in the process he has disembodied affect, locating it outside the subject in the world of technically produced images" (Hansen 2004: xx). As such, Deleuze's movement-image seems more inspired by Bergson's acentred universe of the "aggregate of images," rather than Bergson's more subjectively orientated "centre of indetermination". According to philosopher of new media Mark Hansen

> "the frame in any form cannot be accorded the autonomy Deleuze would give it since its very form (in any concrete deployment) reflects the demands of embodied perception … Beneath any concrete "technical" image or frame lies ... the framing function of the human body qua centre of indetermination (ibid.: 8)."

In this sense, the affective can only arise in the interplay between the image and the player. The affective in the videogame image then refers to those images, or continuous moments within the movement-image, by which the perceived movement from the videogame image passes through the player's body only to be restored to the image by the player's action.

The nature of action: three cases

The movement-image can thus be used to analyse the nature of action in the playing of first and third-person videogames. What follows is a brief exploration of how action relates to both perception and affection in three different videogames by looking at the images they bring forth. In the case of *Doom 3*, the experience of action is bound up in a triad alongside perception and affection in a manner that subordinates the perceptive and affective to the action. Both *Fallout 3* and *Shadow of the Colossus* take on a position towards the experience of action which address the player's reflexive abilities, creating either an understanding of interrelations among the different moments of action, or even an estrangement from the action by shifting the emphasis to the affective or the perception.

According to Deleuze, the action-image dominates what he describes at classical narrative cinema, and given the centrality of action in videogame play, one might be inclined to assume that the videogame image in a similar way is governed by the action-image. Galloway even goes as far to suggest that the action-image exists as the base foundation of the videogame (Galloway 2006: 3).

However, this conception of the nature of action in videogames seems to be founded on a rather crude understanding of both the position of action in the movement-image as well as the concept of the videogame. Galloway here seems to take what Darley describes as the action-oriented videogame as a model for *the* videogame. To Darley, the action-oriented videogames involve "the heightening of sensation, evidenced through the necessity for skill with controls, and the heightened impression of kinaesthesia induced by illusionary participation in acts of spectacular risk and speed" (Darley 2000: 157). The best examples of which are (car)racing games, sports games, and the first- and third-person shooter. And even though this type of videogame has been, and perhaps still is, the most dominant type of videogame, its model of action cannot be taken as a general model for action in videogame play.

Apart from Galloway's action-infused model of the videogame, his conception of the Deleuzian action-image fails to take in account the elements that facilitate the action, namely perception and affection. The nature of action in videogame play can not be reduced to the action-image, as the action-image in Deleuze's understanding is intrinsically connected to both the perception-image and the affection-image (Mukherjee 2008: 232). The point in case however is, that the action-oriented game reduces the realm of the perception- and the affection-image, by which perception and affection appear subordinated to the action. Taking the game of *Doom 3* here as an example, the images of this game are densely populated with a multitude of opponents. In this game, perception is directly subordinated to action not only in the case of the player's actions; once you as a player have sight of your opponents, your opponents will have sight of you and they will not hesitate to attack you, forcing the player in a similar way to couple perception to the response of hostile engagement. The perception in *Doom 3*, then, seems to be characterized by a type of seeing that is already limiting the richness of the visual field in order to focus on the action. This does not mean that the gaze is limited; on the contrary, the player has full agency over the frame, allowing him or her to use the virtual camera in order to frame the image in a way that best allows one to perform the necessary actions. This limitation on perception in *Doom 3* rather extends itself into the spatial representation of the videogame image. The action of unfolds in a limited space

constructed of a maze of corridors and rooms. In this regard, the spatial construction of *Doom 3*, reduces the possibilities for perception to appear in itself, as it redirects and opens up perception directly to the action.

It follows, that if the perception in *Doom 3* flows so directly into action, the interval between perception and action becomes reduced, leaving little room for the affective to occupy the interval. That is to say, there is very little time for the player to register movement on the screen, have it pass through him or her, and to restore the movement on the screen by executing movement. In this regard, *Doom 3* does not primarily play on the affection-image, and as far as affection is called upon, it can be said to play on the body in a primordial sense. The affection-image then seems to take the path of what Deleuze describes as the impulse-image[4]. The impulse-image mediates between affection and action and is constructed of fetishes, or fixations, of Good or Evil (Deleuze 1989: 31). As such, in the context of *Doom 3* it functions to conduct the perception of an alien opponent to foster an embodied yet impulsive reaction through an intuitive yet fast assessment of its threat – or its evilness. In playing *Doom 3* the player is not asked to reflect upon the given situation, dwelling on the affection-image, weighing the different possible actions that the situation might inspire. On the contrary, the player is merely asked to act according to an already conceived logic; shoot, or you will be shot yourself.

The videogame *Fallout 3* can serve as an example of a videogame whose image deploys the elements of the movement-image in a way that allows more room to the interval between perception and action. This is evidenced first by the spatial representation of *Fallout 3*'s videogame image; in contrast to the confined spaces of *Doom 3*, *Fallout 3* presents the open space of a post-apocalyptic waste land, open to perceptive exploration. This space, as a waste land would suspect, is not densely inhabited by non-player-characters, and in these images, perception is not always directly confronted with action. Action, unlike in the action-oriented videogame, is not forced upon the player-character in a narrow confined space, rather, it has to be looked for. In this sense, the tie between perception and action is loosened, and perception emerges in a way that is not already from the viewpoint of a predetermined and unavoidable action. Here too, the player has full agency over the virtual camera, allowing him or her to frame the image with a gaze that is free of immanent action.

This is not to suggest that *Fallout 3* does not contain action, yet the action-image is presented as something that forgoes simply shooting or being shot. *Fallout 3* does present instances where one has to engage the player-character in a fire-fight, but, interestingly, it also contains numerous occasions where the player can choose whether or not to act hostile towards non-player-characters. Moreover, the game presents a multitude of different actions to the player, from collecting a wide array of items, to repairing machines and weaponry, and from engaging in conversation to helping non-player-characters on various quests. At many points in this game, the player is offered an array of different choices to respond to a given situation, by which the need for action separates itself from determining what action is to be taken.

This loosening of the bond between the perception-image and the action-image in *Fallout 3*

4 “Between the perception-image and the others, there is no intermediary, because perception extends by itself into the other images. But, in the other cases, there is necessarily an intermediary which indicates the extension as passage. This is why, in the end, we find ourselves faced with six types of perceptible visible images that we see, not three: *perception-image*, *affection-image*, *impulse-image* (intermediates between affection and action), *action-image*, *reflection-image* (intermediate between action and relation), *relation-image* (Deleuze 1989: 32).

consequently opens up more space for the affection-image to fill the interval. This gap is filled by an intricate web of relations among the different choices and actions made throughout playing the game. For the player-character in *Fallout 3* operates with what can be understood as an adaptive personality; the actions taken will affect the player-character's personality through a series of discrete categories, social skills, technical skills and knowledge, weapon skills, etc. These alterations to the player-character's personality in turn affect the character's abilities to traverse the game; diminished social skills due to being violent towards non-player-characters or impolite in conversation will for instance limit the player's ability to gain information from other non-player character's. On the other hand, by frequently engaging in hostile confrontations of your own choice, the player-character will develop better combat skills and an advanced ability to repair weapons. The actions taken in *Fallout 3* are therefore interconnected, and continue to act upon the possibilities of future actions. In this regard, rather than merely passing through the affective on the level of the impulse-image, the affection-image in *Fallout 3* extents itself to what Deleuze describes as the reflection-image and the relation-image[5], referring to the consequences of the executed actions on the now altered state of the game. In contrast to an action-oriented videogame such as *Doom 3*, the affection-image in *Fallout 3* does not merely pass through the player on the level of impulse, but stays with a player for a extended period of time allowing him or her to contemplate its meaning and relation to the yet to be taken action.

The videogame *Shadow of the Colossus* takes an approach to action which draws into question the logic of action on a more fundamental level. This videogame, while operating within the limits of the movement-image, forefronts the perception-image as to question the correlation between what we see and what we do. *Shadow of the Colossus* is a relatively simple game, about a young man – named Wander – who is set on a quest to slay sixteen giants in order for the gods to bring back to life his love. The player however has to find these giants him- or herself, exploring the game's space on horseback. The action of battling these giants stands in stark contrast to the desolate world the player finds Wander in, presented to the player through action-less, and even near motionless images. No one inhabits this space but birds and other animals, there is – in contrast to *Doom 3* and *Fallout 3* – no action or movement towards the player. Having full agency over the virtual camera while riding through this desolate landscape, the player has all the freedom to look around at will. Yet, in absence of any perceivable action, this perception has little to attach to, becoming, in a sense, cut off from the action-image.

The game's ambiguous stance towards action is brought to its heights, however, in moments of unambiguous action, i.e. the confrontations with the giants. Upon finding a giant, one would perhaps expect to enter automatically into the realm of action, but the giants of *Shadow of the Colossus* are often case not the least bit bothered by, or interested in the player-characters presence. Only when one chooses to attack the giant, will it rise to its defence. In a remarkable way, then, the player has to create its own possibilities for action, as no action will come to him or her of it self. Yet, at the moment of action, something quite out of the ordinary happens to the perception: the subjective perception characteristic of the 3D videogame, which arises from both the subjective manner of framing and the player's ability to control this frame at will, is broken. During the battling of the giants, the player will lose (to a relative degree) the control over the virtual camera's movements and the scale with which it frames the videogame image. If the player, for instance, has the player-character hanging on the back

5 See note 4.

of a giant, and the giant is to make a simple side step, it can occur that this movement by the giant takes the player-character from the frame, when the virtual camera, with some degree of lag trails behind the player-character. These momentary instances in which the player-character swings off-screen, put an added complication to the task at hand. In this sense, perception is not fully subordinated by action in *Shadow of the Colossus*, and even in those parts of the game where one would expect the action-image to dominate the videogame image, perception appears in itself.

This videogame then presents a detachment of perception from action, both in the desolate space that is devoid of action and in the action-packed sequences of the giant battles, which are also permeated by a perception that moves away from action. There appears no visible reason for action, the action in *Shadow of the Colossus* is not a re-action to movement already delivered by the videogame image. The videogame's story also does not provide sufficient motivation for action – sure, Wander wants his love to not be dead, but he only arrives at killing giants because he was told to do so. By fore fronting the perception-image, *Shadow of the Colossus* forces its players to look at the nature of, in this case, unprovoked violent action. It is this question which occupies the affective, for the interval between perception and action is stretched up to the point where perception no longer automatically flows into action. Rather, the abundance of time to look around and see nothing (no movement or action that is), the lethargic attitude of your perceived opponents, and the frustrated perception during the action sequences opens up the question as to what exactly you are doing as a player.

Conclusions

In order to bring into view the relevance of the videogame image in the experience of playing a videogame, the concept of action needs to be united with the physical player. Drawing on the phenomenology of Merleau-Ponty, it becomes possible to appreciate the importance of embodied and enactive perception to the activity of playing a videogame. Action in videogame play is both an embodied and a cognitive activity; it is through the flesh of our bodies that we perceive the images on the screen, and it is through our bodies that we are able to act upon these perceived images.

In the context of digital imagery, the image can no longer be accorded the autonomy Deleuze's ascribes to the cinematic movement-image, and the concept needs to be reworked from a theory centred purely on images to one that focusses on the images and the experience thereof. This means that the elements of perception, affection, and action inherent to the movement-image need to be liberated from the images and need to be considered in relation a phenomenologically constructed notion of the physical player. Possibilities for this liberation can be found in the videogame image's oscillation between visual representation and the control-panel function of the interface-image. It is precisely because the videogame image and the experience of it goes beyond mere representation and visual perception that the videogame image opens up its affective and active facets to the player. The videogame image unites the moment of perception by which the image, through the eyes, enters the body only to be extended into movement again by the player's pushing of the buttons. The videogame image then literally passes through the player, creating a cybernetic loop of movement between and in the player, the game and the computer, and the image.

The employment of a phenomenological infused reading of Deleuze's movement-image in the investigation of the videogame image has not only allowed us the explicate the physical involvement of the player with the videogame image – through the triad of perception, affection, and action – but it also provides us with the means to form a more nuanced understanding of the nature of action within 3D videogame play. If videogames are actions, as Galloway maintains, then it is through the concept of the movement-image that we are able to analyse how the type of action afforded by different 3D videogames is depended on the manner in which the images of these game address the player's perceptive and affective faculties. Taking the videogame image as a point of entry, then, ultimately results in the possibility to formulate a more sophisticated understanding of the nature of action afforded by 3D videogames.

Games

Doom 3. id Software, PC, 2004.
Duke Nukem 3D. 3D Realms, PC, 1996.
Fallout 3. Bethesda Game Studios, PC, 2008.
Hitman: Blood Money. IO Interactive, PC, 2006.
Pro Evolution Soccer 6. Konami, PC, 2006.
Tomb Raider. Core Design, Playstation, 1996.
Shadow of the Colossus. Team Ico, Playstation 2, 2005.
Splintercell: Pandora Tomorrow. Ubisoft Shanghai, PC, 2004.

References

Aarseth, E. (1997). *Cybertext*. Baltimore: John Hopkins University Press.
Aarseth, E. (2001). "Computer Game Studies, Year One". In *Game Studies* July 2001; vol. 1, issue 1, retrieved December 22, 2011, from http://www.gamestudies.org/0101/editorial.html
Aarseth, E. (2004). "Genre Trouble: Narrativism and the Art of Simulation". In *First Person: New Media as Story, Performance, and Game*. Wardrip-Fruin, N. and Harrigan, P. (eds.). Cambridge, Mass.: MIT Press.
Bergson, H. (2002). *Henry Bergson: Key writings*. Ansell-Pearson, K. and Mullarkey, J. (eds.). London: Continuum.
Bogue, R. (2003). *Deleuze on Cinema*. New York: Routledge.
Darley, A. (2000). *Visual Digital Culture*. London: Routledge.
Deleuze, G. (1986 [1983]). *Cinema 1. The Movement-Image*. (trans.) Tomlinson, H., Galeta, R. Minneapolis: University of Minnesota Press.
Deleuze, G. (1989 [1985]). *Cinema 2. The Time-Image*. (trans.) Tomlinson, H. and Galeta, R. Minneapolis: University of Minnesota Press.
Galloway, A. (2006). *Gaming: essays on algorithmic culture*. Minneapolis: University of Minnesota Press.
Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Broadway: Lawrence Erlbaum Publishers.
Hansen, M. (2004). *New Philosophy for New Media*. Cambridge, Mass.: MIT Press.
Hermans, C. (2004). "Trilogy". In *Extensions, The Online Journal of Embodied Technology*, retrieved December 24, 2011, from

http://www.ahk.nl/fileadmin/download/theaterschool/amch/carolien-trilogy-paper.pdf
Jones, M. (2005). “Composing Space: Cinema and Computergaming – The Macro Mise-en-Scene and Spatial Composition” In *Proceedings of the Imaginary Worlds International Symposium*, Sydney, retrieved December 17, 2011, from http://www.dab.uts.edu.au/research/conferences/imaginary-worlds/composing_space.pdf
Jones, M. (2007). “Vanishing point: spatial composition and the virtual camera”. In *Animation*, vol.2, issue 3, pp.225-253.
Manovich, L. (2001). *The Language of New Media*. Cambridge, Mass.: MIT Press.
Marrati, P. (2003). *Gilles Deleuze: cinema and philosophy*. Baltimore: John Hopkins University Press.
Merleau-Ponty, M. (1965[1942]). *The Structure of Behaviour.* (trans.) Fisher, A.L. London: Methuen & Co. Ltd.
Merleau-Ponty, M. (2002[1945]). *The Phenomenology of Perception*. (trans.) Smith, C. London: Routledge.
Mukherjee, S. (2008). “Gameplay in the Zone of Becoming: Locating Action in the Computer Game”. In *Conference Proceedings of the Philosophy of Computer Games 2008*. Günzel, S., Liebe, M. and Mersch, D. (eds.). Potsdam: Potsdam University Press
Nitsche, M. (2008). *Video game spaces*. Cambridge, Mass.: MIT Press.
Noë, A. (2004). *Action in Perception*. Cambridge, Mass.: MIT Press.
Shaw, I. and Warf, B. (2009). “Worlds of affect: virtual geographies of video games”. In *Environment and Planning A 2009*, vol. 41, pp. 1332-1343.
Tavinor, G. (2009). *The Art of Videogames*. Southern Gate: Wiley-Blackwell.

Figure

Figure 1a, 1b, 1c, and 1d. Produced by author.