

Effect of Camera and Object Motion on Visual Load in 3D Games

David Milam, Magy Seif El-Nasr, Dinara Moura, and Lyn Bartram

School of Interactive Arts + Technology (SIAT), Simon Fraser University Surrey, 250 -13450
102 Avenue, Surrey BC V3T 0A3 Canada
{david_milam,magy,dinara,lyn}@sfu.ca

Abstract. 3D video games are a popular form of entertainment that features elaborate visual compositions and settings. Occasionally, players are thrown into situations with a high amount of visual complexity, which may cause players, especially novice players, to misinterpret important game goals or mechanics, which may, in turn, lead to a diminished experience. Previous research investigated visual design, but only in terms of brightness and color contrast, no research investigated attributes of motion, their complexity in regards to visual design, and their effect on the game experience. Informed by cinema and visual perception domains, we embark on a study of 4 motion attributes: flicker, shape, speed, and repetition, and investigate their design within 6 games. We rate these games based on their complexity. We use video coding with a kappa reliability measure to identify these attributes. We then investigate the relationship between the use of these motion attributes and the rated complexity of the visual scene. We present this analysis as a contribution, and design lessons extrapolated based on the analysis.

Keywords: Game Design, Motion, Visual Composition, Visual Perception, User Experience.

1 Introduction

3D video games are popular and top selling games [1]. They feature graphically rich environments, which engulf their users and are designed with an amalgam of challenges including searching for artifacts, wayfinding towards a goal, and visual tracking of moving objects such as enemies. These challenges require players to perceive visual cues, or distinct visual elements relative to their surroundings. Game designers use visual elements to communicate to the player, by nudging, pushing, and pulling the gameplay; this technique offers players a combination of reward, challenge, and punishment in a very subtle form. Therefore, the art of composing these visual elements, which we will call visual design, is extremely important. Communicating gameplay goals to the player is one function of visual design. Designers use visual elements for many different functions, including composing the ambiance and style for the game, which includes composing ambient and atmospheric effects. This *look and feel* is often a defining characteristic of the franchise; it is what makes a game memorable.

One challenge faced by game designers is how to control the visual composition of many elements during moments of high visual load. We use the term *visual load* here to define a measure of complexity of the visual scene. The measure of complexity relies on the attributes of the visual scene (e.g., the number of visual stimuli in the scene) and the cognitive and perceptual processing of relevant and irrelevant visual stimuli. Thus, a high visual load would indicate a very complex environment where participants may find it difficult to differentiate between important visual cues from irrelevant visual elements in the environment, thus leading to a behavioral breakdown [2], [3]. Breakdowns occur when the player misperceives characteristics of the environment, fails to learn a strategy, decision-making step. When left unchecked, breakdowns diminish accessibility and negatively impact player's investment [4].

Previous work on visual design discussed the role of color and lighting techniques [5-8] for instance to enhance emotion and support way finding. Although motion is highly perceptually salient [9], we found no research targeting the use of motion attributes in games. In this paper, we attempt to start work within this area, in the hope of carving the space towards a better understanding of the visual design process within games.

Before games, filmmakers, theatre designers, and visual artists spent many years composing their craft and documenting their techniques. The use of visual composition to direct viewers' eyes using camera, lighting, motion, and color contrast cues is well known within these fields [10-12]. Games follow similar methods, however, games are interactive and thus the ability to decipher the message encoded in the visual cues within the scene is very important. There seem to be a difference between novice and expert players in terms of how they perceive visual scenes [13]. Some scenes that may appear chaotic to a novice player may appear very well composed to an expert player. An understanding of these differences will help formulate design lessons to better develop visual designs that appeal to different groups of gamers.

In this paper, we investigate how games organize motion attributes. Specifically we ask 2 research questions: in what way do game scenes rated with low visual load constrain and organize motion relative to the camera? In what way do game scenes rated as high visual load organize motion attributes? We believe this is the first step towards developing a theory around dynamic adjustment of visual cues, similar to the work done previously on lighting [5], [14].

This paper investigates visual complexity based on theories of visual perception including the preattentive (bottom-up) processing of salient features [15], similarity [16], [17], and visual search [18] (bottom-up and top down). In particular, we video coded the occurrences of 4 motion attributes: flicker, shape, speed, and repetition, and investigate their design within 6 game scenes selected for their variety in terms of their visual load.

The organization of this paper is as follows. We first discuss theoretical foundations that we use to analyze the motion attributes within the visual scenes, in particular: borrowing from visual perception and attention theories. We follow this discussion by the methods we used to collect and code attributes in 6 game scenes. We then discuss results of the analysis and implications of our findings formulating design lessons.

2 Theoretical Foundation and Previous Work

2.1 Movement in Visual Composition

Ward [11] defined visual composition in film or television as an “invisible technique”, because often several elements are used for a specific function, such as emphasizing the subject, rather than reveal the specifics of the visual elements themselves. This requires understanding human perception and the manipulation of visual design elements to engage the viewer’s attention. For example, the focal point of the scene must solicit attention. Designers often negotiate multiple elements, such as camera position and subject movement to create a design that reflects the focus of attention desired.

In addition, Block [9] discussed several visual compositional elements of films, including space, tone, line, shape, color, movement, and rhythm. He stated that viewers’ attentions are usually drawn to moving objects first followed by the brightest object. He also identified 4 motion attributes that influence the visual composition of a film scene: direction, quality, scale, and speed of motion.

The games industry frequently alludes to the importance of composition through interface design [1] and environment design [4], [19]. The interface is often used to emphasize the execution of explicit actions, while visual cues in the environment concentrate on more subtle design messages through channeling attention and direction of movement.

Several researchers investigated the effect of such environmental cues on path choices [7], [8]. They found that participants chose the path with a moving element, when all else is static. Other studies focused only on task efficiency with low level perceptual cues embedded into the environment, specifically color and brightness contrast, and texture along with texture [6], [20]. For [6], participants were dropped into a maze and their paths recorded to determine navigation preferences based on the low level cues. The analysis showed that low level cues alone are not reliable as most participants did not notice and did not follow the cues. The study by [20] considered both low level cues and top down goals to focus attention such as collection or finding an exit necessary to advance. The authors deduce that top-down visual features tied to goals control players’ attention more than bottom-up visual features. Motion remains an under studied phenomenon especially during moments of high visual load.

2.2 Motion in Visual Perception

While no research work within games investigated the use of motion as a visual design element, most of the research on motion research can be found in visual perception and attention domains. Researchers found that motion has many expressive perceptual properties particularly in the periphery, where other features such as color change are imperceptible [21]. Some research has been devoted to identifying motion properties including speed, shape (the path a motion follows), phase, flicker (defined as discrete flashes perceived by the viewer), and smoothness were emphasized as critical motion attributes [9], [22]. Of these, shape – particularly circular motions – has been shown to be particularly important, especially where communicating affect is concerned [9], [23].

There are many prominent theories of visual attention such as the saliency and similarity of low-level features. Treisman [15] introduced the saliency map as a combination of low level features (stimulus-driven) that are rapidly perceived preattentively and without conscious thought. Additionally, Wolf's guided search theory [18] applies top-down and bottom up approaches working in parallel. Guided search can assign additional attentional priority to low-level features given a cognitive goal or task. These theories are similar to Ware's [24] psychology of visual thinking, whereby our attention is continuously biased in order to accomplish actions and complete tasks. The design implication here is to use clear features with distinct processing channels, as they are less cognitively demanding to perceive.

Another popular approach posed by Duncan and Humphrey is a similarity theory of attention [16], which extends from Gestalt psychology [17]. Similarity considers how easily goal-relevant and goal-irrelevant stimuli can be distinguished. Search efficiency and task time is low when the stimulus is different and high when the differences are subtle. In a related study, Kingstone et al. [25] applied similarity theory to analyze the coherence of motion in a visual search task. They found motion can inhibit or excite search efficiency. They predicted that: 1) search is easier when there are fewer motion groups, 2) it is easier to inhibit rather than excite goal-relevant stimuli, and 3) search gets easier as stimuli move together.

Previous works in visual attention and games have received scientific research in experimental and video game contexts. For example, researchers [26], [27] tested participant reaction time to rapid visual stimuli located on the screen center and periphery. Their results showed that gamers with more experience had decreased reaction times. However, since they didn't test these results within a commercial game, it is hard to confirm the ecological validity of these results.

3 Method

3.1 Data Collection Procedure

To investigate how designers varied the motion attributes and the relation between motion attributes and visual load, we collected 2 video clips per game. We selected 6 games for this investigation with varied visual complexity. We selected these games based on the following criteria: (a) recently produced (2009 or later), (b) received high sales, and (c) span the 3rd person action-adventure and role playing genres. All games selected received aggregated reviews from metacritic.com between 70-100, which are "generally favorable" and received "universal acclaim". The two videos we gathered were from Youtube walkthroughs. Thus, they have been viewed thousands of times.

The selection of each video clip contains gameplay with a high amount of visual load due to motion. Because the density of motion is very high, we only chose short clips, thus each clip is approximately 20 seconds in duration. Selections are from the game demo or first instance where high load occurs within the first hour of play. All clips contained one or more tasks for the player to complete, such as using weapons or tools, selecting or targeting objects, collection, managing resources with assistance from the user interface, and maneuvering through dynamic 3D settings.

Table 1. Data collection from 6 Games with abbreviations

Game Name	Abbreviation
Assassin's Creed Brotherhood	AC2
God of War III	GOD3
Prototype	PRO
Ratchet & Clank Future: A Crack in Time	RATCH
Fable III	FA3
Prince of Persia: The Forgotten Sands	POP

The selected clips were then shown to 3 additional researchers participating in the project. Researchers were asked to qualitatively rank each video game from lowest to highest amount of visual load on a 5-point scale. All researchers are familiar with attributes of motion, discussed in the next section, as they came from scientific visualization or performance arts fields.

3.2 Coding for Motion and Camera Attributes

Although all games are in 3D, each clip constrains the players view by means of the camera. For our study we tagged each clip with one of 3 camera configurations including 1) stationary, 2) follow, or 3) free. A *stationary camera* is stuck to a specific 3D position and is constrained to a specific direction. There may be some camera wiggle room for the player or camera, but both are in a relatively stationary position. A *follow camera* trails behind players moving forward or sideways along a predefined linear path. A *free camera* has neither stationary nor following constraints as the player has control to look and move in any direction. Figure 2 identifies the camera setting used for each clip. Clips from FA3 were stationary, GOD3 and RATCH used follow cameras, and both clips from PROTO used only a free camera. In addition to the camera constraints we also annotated instances of camera shake, occlusion, and full screen flashes.

Table 2. 3 camera configurations for each clip (C1 or C2), from each game

Stationary	Follow	Free
POP C2	POP C1	ACB C1
FA3 C1 & C2	RATCH C1 & C2	PROTO C1 & C2
ACB C2	GOD3 C1 & C2	

Informed by cinema, visual attention, and perception literature, we identify four motion attributes: flicker (flashing), shape, speed, and repetition. The parameters are shown in table 3. Flicker refers to flashing or pulsing (once, repeating), the parameters for shape are (linear, circular, expansion, or expansion with contraction), speed (slow - fast), and repetition. Repetition refers to regularity or rhythmic patterns of motion such as harmonic motion. Repetition has a mathematical regularity, such as

wave, sinusoidal, or oscillating forms. Attributes were annotated for all game objects, both UI interface objects (controls stuck to the screen) and non UI objects/effects.

Two raters coded for these attributes in the video clips selected. We used the kappa measure as a statistical measure of inter-rater reliability. Our kappa scores shows almost perfect agreement in all 4 attributes for the flicker, shape, speed, and repetition, with a kappa value of 0.898, 0.847, 0.9, 0.967, respectively. For each clip we compiled a list of UI or game objects that contained at least one attribute of motion that “popped out” or contained high saliency. On average 5 objects are identified per clip (min = 4, max = 8) with 68 total objects coded for all clips. Each object contained up to four attributes of motion.

Table 3. 3 camera configurations for each clip (C1 or C2), from each game

Attribute	Parameters
Flicker (flashing)	none, once, repeating
Shape	none, linear, circular, expand, expansion/contraction
Speed	none, slow - fast
Repetition	none, has repetition

4 Results

We organize results in relationship to the original research questions. The games are first ranked identifying differences in the use of motion attributes for games rated with high and low visual load. We then investigate in depth the motion attributes per clip given the camera constraints comparing games rated as high or low visually loaded.

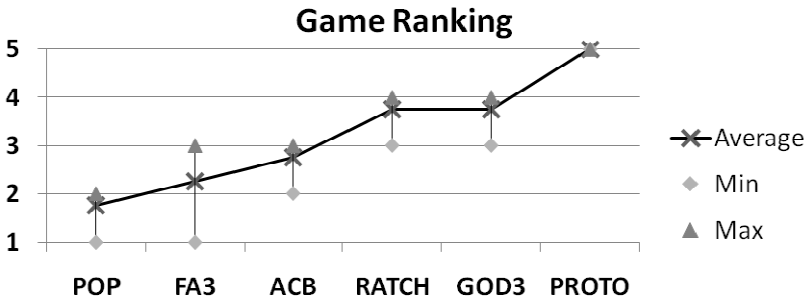


Fig. 1. Ranking of 6 games according to perceived visual load (1=least, 5=most)

Four participants ranked the 6 games from least to most visual load on a 5-point likert scale. Results are shown in figure 1. Prince of Persia was rated lowest while Prototype was rated highest. Fable3 and Assassin’s Creed Brotherhood were rated neutral or below while Ratchet and Clank and GOD3 were rated neutral or above. From this ranking we can investigate the attributes further and their relationship to the

ranking. While it is easy to determine Prototype has the highest load due to the free orient camera, Assassin's Creed also used a similar camera and was ranked lower than games that constrained the camera. Therefore, there is more involved than just the type of camera used.

4.1 Influence of Camera on Motion Attributes

Before discussing the attributes of motion specifically, this section discusses 2 camera techniques to organize attributes of motion. The techniques include constraining the camera to a fixed position and view or following a predefined path of travel. From these constraints we identify focal points and visual regions as organizational strategies to layer motion attributes. Focal points anchor or embellish motion to a central point of attention that is separate from the background. From this focal point multiple attributes of motion can be layered. Clips from ACB, FA3, and POP used a stationary camera. For example, fast motion anchored to a point occurred either over a static or slow moving background. In addition, the focal point duration is preserved for the whole clip or for several seconds. In figure 2 from FA3, circling bats, expansion-contraction magic particles, linear spell cast, and explosions are all anchored to the player. Yellow lines are added to the screenshot and a diagram is added for clarity. The same focus occurred in POP clip2 and FA3 clip 1.



Fig. 2. FA3 C1 circular, linear, expansion and contraction shape is tied to the player

Second, the use of a follow camera along a predefined path of travel helps separate distinct visual regions. This enabled POP, RATCH and GOD3 to excite or recede regions into the foreground or background. A simple example is in POP clip 1, which is shown in figure 3, where both the player and camera are climbing from right to left on a wall. Far below the player are dozens of army soldiers moving *slowly* while a *linear* shaped and *fast* speed fireball approaches from left to right. The effect is very dramatic as it highlights a focal point of interest over a background of slow activity. This technique is also used in clips from RATCH and GOD3 where fast, slow, and repeating motion was used to distinguish between active and inactive regions. Both games also used focal points centered on the player avatar, to communicate resource collection and combat actions. For instance, figure 4 shows a screenshot from RATCH clip 2 with dozens of excited objects above the horizon (ground is static), slow moving and repeating circular trajectories. Players must advance underneath this active area.



Fig. 3. Screenshot (left) and motion diagram (right) for POP clip 1

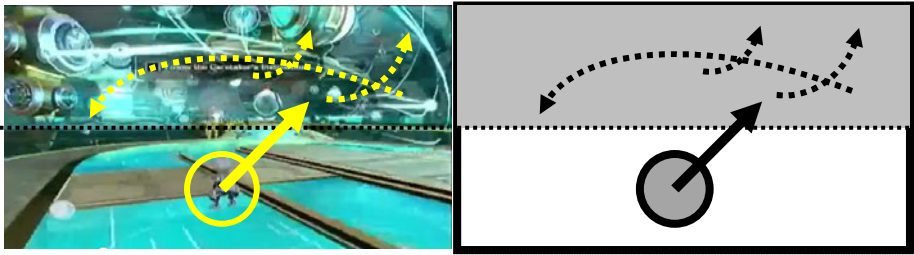


Fig. 4. RATCH, the player moves on a path between a static floor and beneath active objects

The free orienting camera, unlike the stationary and follow camera are without a sustained focal point or visual region. Games that used a free camera, ACB clip 1 and PROTO have a high visual load rating for multiple reasons. First, in every passing second there are changes in views, and thus there are constant shifts in focus from one object or region to another. Second, foci identified through expanding-contracting shapes and fast speeding objects are spread over the entire screen area, and superimposed without a clear visual hierarchy. For example, PROTO displays multiple linear trajectories at once, including 1) from enemy crossfire originating from the screen bottom, 2) a second enemy position falling backward and away towards screen right, and 3) a third enemy on the top of the screen running away from the player. At this moment there is no dominant pathway or object to orient towards. In this case fast expand-contract (explosion) and fast moving enemies occur without regularity or rhythm, which also contributes to the high load ranking.

4.2 Analysis of Motion Attributes

Next we identify four key differences in the use of motion attributes for games rated with high and low visual load. Due to limitations of space, the analysis is brief.

For the speed attribute, use of a free orienting camera with several fast moving objects has high visual load. Both clips of PROTO have many objects all with fast speed while this was not a problem in ACB with only 1 fast object. The use of a stationary camera with fast moving objects has low visual load. This was the case for FA3 and POP. It should be mentioned that in these games, fast moving objects also

converged around the player avatar as a central focal point of interest, which gave the visual scene better coherence and focus as well. Games that used a follow camera, RATCH and GOD3, had contrasting fast and slow moving objects as a means to highlight both a focal point, excite, or recede a visual region. The use of both techniques simultaneously can explain the higher rated amount of visual load.

For Flicker, we did not find the amount of flashing to have an effect on visual load ratings. 11 of 17 instances of flicker were small in scale and occurred on the interface with FA3, GOD3, and ACB flashing both the interface and an embedded object simultaneously. The PROTO clips contained multiple flashing elements, however, they were not constrained to a focal point or a visual region. Unconstrained flashing with a free orienting camera can explain the higher rated amount of visual load.

The motion repetition attribute is higher in games using a follow or stationary camera. GOD3 and RATCH contained both a high amount of motion-repetition and speed variation, which contributed to their higher load ratings. These 2 games in particular had the greatest variety of motion shape attributes, but otherwise the amount or type of shape alone does not appear to be a factor separating high and low load games.

5 Implications of Visual Load from Camera and Object Motion

Player retention [4], preserving accessibility [28], and maintaining motivation [29], [30] are frequently discussed topics in the games industry. This paper considers moments of high visual load as a critical moment where players may stop playing due to a misinterpretation of visual layering or frustration. We present 3 lessons as a result of our analysis.

1. Use motion to anchor a focal point or define a region in parallel with a constrained camera. Adding constraints to the camera will help guide attention and behavior for longer time durations. Attributes of motion can be layered as long as they are anchored to a visual focal point using a stationary camera. Follow cameras can also establish a strong visual region through the use of motion repetition. Free orienting cameras not only require additional motor control, but important focal points or regions on the screen can easily become lost with the wrong point of view. Although players have more freedom and control, there is an enormous compositional cost in reducing the visual coherence.
2. Use contrasting attributes of motion to strike a balance. Games look stale if too many elements are static and can easily overwhelm when everything moves at once. Consider layering or creating hierarchy of motion by combining flicker, shape, speed, and rhythmic repetition. We found many examples where motion repetition with a variable speed and shape can excite or recede a visual region.
3. Prioritize repeating, harmonic or rhythmic motion to visually ignore or excite objects. Our visual system is wired to detect visual patterns such as similarity or coherent grouping. Use this knowledge to visually ignore or excite visual focus and promote region separation. For games with a high amount of visual load spread across a large area, repetition or harmonic motion should be used.

6 Conclusions and Future Work

Our primary contribution was to analyze the relationship between rated visual load, using a camera classifying scheme, and motion attributes, specifically flicker, shape, speed, and repetition. We found that camera configurations, specifically a stationary and follow camera, play a large role in establishing a focal point and visually separating regions. We present the analysis and resulting design lessons as the contributions for this study in the hope of building the foundation for the study of motion in visual design of games.

The study discussed has several limitations. In particular, we didn't discuss camera effects such as shaking, blurring, or full screen flashing. Low visual load situations were not included because as the previous research suggests, players can easily identify single moving targets when the rest of the scene is static. Additionally, we do not consider how designers habituate or train players to learn a visual language. Finally, stylistic differences between games have different constraints on motion. For example photorealistic games such as ACB and PROTO contain believable renditions of actions while stylistic games RATCH and GOD have amplified and exaggerated effects.

In future studies, we aim to apply a subset of this framework into a tool defined by the motion attributes and follow this by a participant study. We hope to identify more specifically the relationship between motion attributes and breakdowns in an experimental study where such attributes can be manipulated. We believe this will allow us to more objectively study the effect of motion on breakdowns and players' experience.

Acknowledgments. The Expressive Motion project is funded by Canada NSERC. Special thanks to Diego Maranan and Matt Lockyer who provided group rating assistance in this study.

References

- [1] Adams, E.: *Fundamentals of Game Design*, 2nd edn. New Riders, Berkeley (2009)
- [2] Ryan, W., Siegel, M.A.: *Evaluating Interactive Entertainment using Breakdown: Understanding Embodied Learning in Video Games*. In: *Breaking New Ground: Innovation in Games, Play, Practice and Theory*, West London, UK (2009)
- [3] Seif El-Nasr, M., Milam, D., Aghebeigi, B., Canossa, A., Niedenthal, S., Maygoli, T.: *Game Metrics based on Error Measures: A deeper look into your players' behaviors*. In: *Game Metrics: Maximizing the Value of Player Data*, tbd (2011) (forthcoming)
- [4] Smith, R.: *Helping Your Players Feel Smart: Puzzles as User Interface* (2009)
- [5] Seif El-Nasr, M.: *Intelligent Lighting for Game Environments*. *Journal of Game Development* 1(2) (2005)
- [6] Samarinas, A.: *Illuminating Minotaur's Lair: Light Design and Navigation in Gameworlds*. MSc, IT University of Copenhagen (2009)
- [7] Hoeg, T.: *The Invisible Hand: Using Level Design Elements to Manipulate Player Choice*. In: *Masters of Interactive Technology in Digital Game Development with a Specialization in Level Design*. Guildhall at Southern Methodist University (2008)

- [8] Alotto, B.: How Level Designers Affect Player Pathing Decisions: Player Manipulation through Level Design. In: *Masters of Interactive Technology in Digital Game Development with a Specialization in Level Design*. Guildhall at Southern Methodist University (2007)
- [9] Bartram, L., Ware, C.: Filtering and Brushing with Motion. *Information Visualization* 1(1), 66–79 (2002)
- [10] Block, B.: *The Visual Story: Seeing the Structure of Film, TV, and New Media*. Focal Press, Boston (2001)
- [11] Ward, P.: *Picture Composition*, 2nd edn. Focal Press, San Francisco (2003)
- [12] Solso, R.: *Cognition and the Visual Arts*. MIT Press, Cambridge (1993)
- [13] Trefry, G.: *Casual Game Design: Designing Play for the Gamer in All of Us*. Morgan Kaufmann, Burlington (2010)
- [14] Seif El-Nasr, M., Vasilakos, T., Rao, C., Joseph, Z.: Dynamic Intelligent Lighting for Directing Visual Attention in Interactive 3D Scenes. *IEEE Transactions on Computational Intelligence and AI in Games* 1(2) (2009)
- [15] Treisman, A.M., Gelade, G.: A feature-integration theory of attention. *Cognitive Psychology* 12(1), 97–136 (1980)
- [16] Duncan, J., Humphreys, G.W.: Visual search and stimulus similarity. *Psychological Review* 96(3), 433–458 (1989)
- [17] Wertheimer, M.: Untersuchen zu lehre von der Gestalt [Laws of Organization in Perceptual Forms]. *Psychologische Forschung [Psychological Research]* 4(1), 301–350 (1923)
- [18] Wolf, J.: Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review* 1(2), 202–238 (1994)
- [19] Rogers, S.: *Everything I Learned About Level Design I Learned from Disneyland* (2009)
- [20] Seif El-Nasr, M., Yan, S.: Visual Attention in 3D Video Games. In: *Presented at the SIGCHI International Conference on Advances in Computer Entertainment, Hollywood* (2006)
- [21] Bartram, L., Ware, C., Calvert, T.: Moticons: detection, distraction and task. *International Journal of Human-Computer Studies* 58(5), 515–545 (2003)
- [22] Huber, D., Healey, C.: *Visualizing Data with Motion*, pp. 527–534 (2005)
- [23] Bartram, L., Nakatani, A.: What makes motion meaningful? Affective properties of abstract motion (2010)
- [24] Ware, C.: *Visual thinking for design*. Morgan Kaufmann Publishers, San Francisco (2008)
- [25] Kingstone, A., Bischof, W.F.: Perceptual grouping and motion coherence in visual search. *Psychological Science* 10(2), 151–156 (1999)
- [26] Castel, A., Pratt, J., Drummond, E.: The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica* 119(2), 217–230 (2005)
- [27] Green, S., Bavelier, D.: Effect of Action Video Games on the Spatial Distribution of Visuospatial Attention. *Journal of Experimental Psychology: Human Perception and Performance* 32(6), 1465–1478
- [28] Desurvire, H., Caplan, M., Toth, J.A.: *Using heuristics to evaluate the playability of games*, Vienna (2004)
- [29] Klimmt, C., Hartmann, T.: Effectance, self-efficacy, and the motivation to play video games. In: Vorderer, P., Bryant, J. (eds.) *Playing Video Games: Motives, Responses, and Consequences*, pp. 132–145. Lawrence Erlbaum Associates, Mahwah (2006)
- [30] Phillips, B.: *Staying Power: Rethinking Feedback to Keep Players in the Game*. Gamasutra Online Magazine (October 2009)