

# Martial Arts in Artificial Reality

**Perttu Hämäläinen, Tommi Ilmonen**  
Helsinki University of Technology  
P.O.Box 5400, FIN-02015 HUT,  
Finland  
{pjhamala,tilmonen}@tml.hut.fi

**Johanna Höysniemi**  
University of Tampere  
Kanslerinrinne 1, FIN-  
33014 UTA, Finland  
johanna@cs.uta.fi

**Mikko Lindholm, Ari Nykänen**  
Animaatiokone Industries Co-op  
Sakarinkatu 3 B 52, FIN-00520  
Helsinki, Finland  
{mikko,ari}@animaatiokone.net

## ABSTRACT

This paper presents Kick Ass Kung-Fu, a martial arts game installation where the player fights virtual enemies with kicks and punches as well as acrobatic moves such as cartwheels. Using real-time image processing and computer vision, the video image of the user is embedded inside 3D graphics. Compared to previous work, our system uses a profile view and two displays, which allows an improved view of many martial arts techniques. We also explore exaggerated motion and dynamic slow-motion effects to transform the aesthetic of kung-fu movies into an interactive, embodied experience. The system is described and analyzed based on results from testing the game in a theater, in a television show, and in a user study with 46 martial arts practitioners.

## Author Keywords

Martial arts, kung-fu, artificial reality, computer vision

## ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces---input devices and strategies, interaction styles.

## INTRODUCTION

Mirrors and video analysis are used in many sports to spot errors in pose and motion. In research literature, several computer assisted motion and biomechanics analysis systems are described. Various approaches include user-assisted video analysis, tracking devices, and computer vision [5,16,20,26]. Previously, we have investigated the use of interactive video mirrors for real-time motion analysis [13], using a setup shown in Figure 1. The basic idea was to combine the benefits of mirrors and video in repeated performing and evaluation of acrobatic and martial arts moves, for example, by delaying the camera view by a few seconds so that you can perform a spin kick and see it

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2005, April 2-7, 2005, Portland, Oregon, USA.  
Copyright 2005 ACM 1-58113-998-5/05/0004...\$5.00.

repeated on the screen immediately, in contrast to a traditional mirror where you cannot see yourself when spinning around.

This paper continues our previous work to motivate training through playful entertainment. We describe the design and evaluation of Kick Ass Kung-Fu, a martial arts game installation where you fight virtual enemies with kicks, punches, and acrobatic moves such as cartwheels. The game is shown in action in Figures 2-3.

With real-time image processing and computer vision, the video image of the user is embedded inside 3D graphics on a virtual playfield facing virtual opponents. Your movements are exaggerated so that you can dodge your opponent's attacks by jumping five meters in the air and landing behind his back. Using the dual projected screens, one at each end of the playfield, you can also continue by counter-attacking the enemy from behind.

Compared to previous work, we describe and evaluate an artificial reality system with two screens and the user shown in profile view with the camera on the side. Although users may need some time adapting to the setup, we show that it works well and we hope that our study encourages others to try out new display and camera configurations.

Movie style exaggerated motion and slow-motion have also



**Figure 1. A video mirror test setup, showing camera (1), camera view projected on a screen (2), user (3) and test instructor (4).**



**Figure 2.** Kick Ass Kung-Fu on stage in a theater. The player moves on a cushioned mat, facing either one of the projected screens. The player sees himself or herself on the screen in a profile view familiar from traditional martial arts games.



**Figure 3.** A screenshot of the game, showing the player on the left, embedded in the game with a drop shadow.

not been researched in the context of embodied interaction, although during the past few years, they have become common in traditional video games such as *Max Payne* and *Prince of Persia Sands of Time*.

In addition to the conceptual developments, we present the results of a user study with 46 martial arts practitioners, discussing the use of the game as part of training and the navigation and mapping problems related to the profile view.

## RELATED WORK

Interactive video mirrors and processed video of the user have been used in games and art installations starting from *VideoPlace* by Krueger *et al.* [17], where the two-dimensional video image of the user interacts with computer generated animated characters. Krueger calls the approach *artificial reality*, but it can also be considered as augmented reality, especially if the background of the user

is not removed and graphics are only overlaid in the camera view. There are also commercial applications of the approach, such as the *Eye-Toy* camera and games for the PlayStation 2 [2]. The MIT Media Lab *Alive* system is an example of a more sophisticated, 3D interactive video mirror where you can interact with computer generated characters with gestures [28].

There are many examples of previous human-computer interfaces that require physical effort. In the wake of Konami's *Dance Dance Revolution* released in 1998 [1], arcades all over the world feature dancing games where you have to perform combinations of steps in rhythm with music. Ishii *et al.* presented an "athletic-tangible interface", a ping-pong table that augments the sport with dynamic graphics and sound [12]. Mueller *et al.* discuss social aspects of computerized sports over distance, based on a game setup with a regular soccer ball and life-size videoconference screen [18]. In the context of martial arts and computers, Chua *et al.* describe a wireless virtual reality system with a head-mounted display for tai chi training [8]. Chi *et al.* present a wearable sensor system that registers impacts in taekwondo competitions [7].

In a broader perspective, our work is related to research on perceptive user interfaces (see, e.g., Crowley *et al.* [9] or Wren *et al.* [28]). In vision based games, an alternative to using the video input as an avatar is to use a computer generated avatar, such as in the *QuiQui's Giant Bounce* game, where a dragon flies when you flap your hands and breathes fire when you shout [14].

## SYSTEM DESIGN

Figure 2 shows a typical *Kick Ass Kung-Fu* setup. You move on a 5x1m cushioned playfield, facing either of the two screens. The screens can be tilted slightly to provide a better view for the audience. The playfield is large enough for you to perform a combination of three kicks or a combination of a cartwheel and a kick. In general, the larger the playfield, the better, but the installation space soon becomes a problem.

### The Profile View

As shown in Figure 3, you see yourself in a third-person profile view that is traditional in martial arts games. The visual design extends previous artificial reality systems with a 3D look and drop shadows generated from the user's video image. The shadows are important for a 3D feel and without them it is impossible to tell whether you are on the ground or in the air.

Both screens show the same view, except that one of the screens may be mirrored horizontally, the effects of which are discussed later in this paper. This is not a CAVE-style multiple screen virtual reality setup, where each screen shows a different part of a first-person view [10]. The two screens are used to let you fight multiple enemies that attack you from both sides.

There's a 90 degree rotation between the world and screen coordinates, which can cause problems similar to when you try to cut your hair using a mirror so that you need to remap the directions and visual feedback. In our earlier video mirror study, people sometimes reached to their right when reaching for onscreen user-interface widgets that actually were in front of them [13]. When designing the system, the first idea was to have only one screen and the user facing a camera placed under the screen. However, the approach was abandoned since the authors knew from their own martial arts experience that your technique suffers when practicing kicks and punches if you are not facing in the direction of your opponent. The profile view shows most kicking techniques optimally so that you can see how your legs extend and how far you reach.

### Sound Effects in World Coordinates

In relation to the profile view, game sound effects are localized in world coordinates instead of screen coordinates. Speakers are placed at each end of the playfield and enemy sounds come from the direction the user should move to reach the enemy. This was requested by a martial arts teacher when testing a Kick Ass Kung-Fu prototype so that the sounds were rendered mostly for the audience with loudspeakers near the ceiling.

### Education vs. Fun: Exaggerated Motion

Our design goal was to create a physically interactive game that would be both fun and useful as an augmenting form of martial arts training. Combining education and fun is generally not easy, but it has been noted that you can learn from realistic game elements, e.g., the weapon systems of war games, even if the game is designed purely for fun [4]. In our case, an important concept is *user interface realism* – learning can be facilitated with a realistic interface, such as the controls and displays of a flight simulator. The user interface of Kick Ass Kung-Fu is realistic in that you actually perform martial arts moves and get visual feedback of your performance.

On the other hand, Oblinger talks about a tradeoff between realism and fun [19]. A major attraction of computer games is that you can try out things not possible in the real world. The same applies to action movies, which would be quite dull if there were no unrealistic stunts.

Exaggerated jumps and slow-motion mid-air acrobatics are an important aspect of the aesthetic of martial arts movies. We wanted Kick Ass Kung-Fu to provide an embodied movie-style martial arts experience, enabled by real-time computer vision and image processing technology. Similar to movie post-production, the user is embedded in a virtual set. The difference to movies is that motion is exaggerated only visually instead of the user hanging from wires like kung-fu actors, which is dangerous without proper training.

### Dynamic Slow Motion

In our previous work with video mirrors, we found slow motion playback to be useful in inspecting martial arts

techniques. Slow motion is also used in kung-fu movies to vary the rhythm of the action and let the viewer appreciate the acrobatic skills of the actors. On the other hand, movies also use accelerated shots, for example to make a series of punches appear more lethal.

In a traditional game played with a gamepad, slow-motion is not problematic since you can press buttons to control the avatar even if it moves at different speeds. In a physically interactive system with a one-to-one mapping between the user and the avatar, slowing down the avatar makes it go out of sync with the user. Kick Ass Kung-Fu employs slow motion dynamically so that when you jump high enough, the camera feed is slowed down, and when you land, the feed is fast forwarded back to real time. You can first fight with normal speed on the ground, then perform a flamboyant slow-motion jump kick and continue with a series of punches that get shown at faster than normal speed.

### Technology

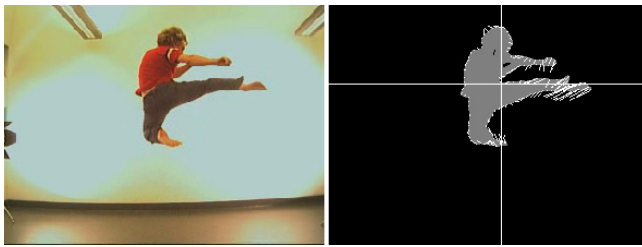
The game runs on a 2.8GHz Pentium 4 laptop computer equipped with a USB webcam. The software uses OpenGL for graphics and the OpenCV library for computer vision and image processing [3].

#### *Computer vision: Background Subtraction and Optical Flow*

Computer vision was selected as the technology instead of floor sensors and wearable devices because it enables full-body unencumbered movements with low hardware cost, there are no breakable parts, and it is easy to switch the player and enable multiplayer gaming. We use a combination of background subtraction and optical flow computation, which enables any number of players to collaboratively fight the virtual enemies. You can also wield practically any weapon.

The player is transferred into the game using background subtraction, of which a good review is given by Toyama *et al.* [25]. We use a basic form of background subtraction. A sample of the static background is captured when the game starts. The pixelwise color difference between each video frame and the background is computed and pixels with large enough differences are considered as part of the user. An OpenGL texture is generated of each frame with other than user pixels as transparent. To render the player inside the 3D graphics, the texture is mapped onto a rectangle. The drop shadows are easy to add by drawing a semitransparent black shadow rectangle with the same texture.

Collisions are detected between enemies and the outline pixels of the user. The velocity of the outline pixels is estimated using the OpenCV implementation of pyramidal Lucas-Kanade optical flow. A hit is registered if a colliding outline pixel moves fast enough, so that damage is caused only by actual attack techniques and not by just touching the enemy. Note that the user interface is pseudo-3D in that the optical flow only registers 2D motion and the user is treated as a 2D plane within the 3D scene.



**Figure 4. Camera view (left) and computer vision debug view of a jump kick. The crosshair shows the mass center of the user pixels and the lines show the optical flow computed.**

It is practically impossible to measure how much force there actually is behind a punch, but measuring speed is more realistic than, e.g., the Eye-Toy Kung Foo game, where you only need to brush the enemies with your fingers to cause the same amount of damage as with a full punch. The enemies can easily be made to fly in the direction of the attack so that you can first kick an enemy from below, launching it up in the air, and continue with another technique as the enemy falls back down. The velocity of the mass center of the user pixels is subtracted from the optical flow so that if you just walk and bump against your opponent, the effective strike force is zero. The optical flow and mass center are visualized in Figure 4.

In order to not register hits constantly when moving your hands or feet inside enemy area, the corresponding outline pixels are found in the previous frame and a hit is registered only if they were not colliding already.

#### *Horizontal Motion: Simple Scaling*

Horizontal motion is exaggerated simply by scaling the motion of the user's mass center. The camera texture is placed at

$$x = (c_x - 1)x_m w, \quad (1)$$

where  $x_m$  is the horizontal position of the user's mass center,  $w$  is the width of the texture and  $c_x > 1$  is the scaling factor. If  $c_x = 1$ , the texture does not move horizontally so that the user's motion in the camera view is shown as such.

In addition to making motion faster, the scaling has the benefit that the virtual playfield can be larger than the real one. However, too much boost makes it difficult to estimate the reach of your attacks. The scaling also makes the avatar appear to slide on the ground when walking or kicking. Fortunately, the sliding becomes clearly visible only with high scale values so that slightly exaggerated motion still looks natural.

While being intuitive, this kind of direct mapping between the user's and the avatar's horizontal position prevents the avatar from reacting to enemy hits. If the avatar was flown back by punches, the mapping would break and the user would eventually end up walking out of the camera view just trying to maintain a fixed position on the virtual playfield. The differences between virtual and real worlds are a problematic issue when designing physical user

interfaces, as noted previously in a case where the avatar flies so that you wave your hands and bend your body in the direction you want to go [14]. Since the user's horizontal location was not mapped to the avatar, the players could unintentionally move out of the camera view.

#### *Vertical Motion: A Hybrid Velocity/Acceleration Mapping*

Exaggerating vertical motion was not found as trivial as horizontal motion. The main reason for this is that in the single camera 2D motion analysis, vertical motion is ambiguous with motion towards or away from the camera. If the user steps farther away, he or she is perceived as moving upwards. Thus, simply scaling vertical motion similar to horizontal motion can lead to the avatar floating in the air. Multiple cameras could be used to estimate the distance and height of user pixels more accurately, but we wanted to keep the hardware setup as simple as possible.

We first tried a straightforward system where the user's vertical position was obtained as the scaled vertical position of the mass center, from which the ground level was subtracted, estimated as a moving average of the vertical position. However, the motion did not appear natural, since you could pump yourself up and down, for example, by lifting your knees to your chest when jumping. Other reference points, such as head and feet were also tried, but similar anomalies still manifested, for example, when doing cartwheels.

In games, natural motion is often created by simulating physics, for example, the effects of gravity on a jumping character. In its most basic form, this can be done using Euler's method to solve the differential equation  $a_y = dv_y/dt$ , that is, by replacing  $dt$  with the simulation timestep and  $dv_y$  with the change in vertical velocity at each simulation step [23]. If no external forces are applied, acceleration  $a_y$  equals gravity. At the beginning of a jump,  $v_y$  is initialized to a suitable value, after which gravity takes care of bringing the character back down.

In our system, the velocity and acceleration of the mass center of the user pixels is measured at each video input frame. Basically, the avatar can be made to jump naturally by applying the measured acceleration to it if the user is launching upwards. If the user is falling down or just standing, gravity is applied so that the avatar falls down until it touches the ground. However, the measured acceleration is sensitive to noise and errors due to the camera driver dropping frames or the user moving away from the camera. When testing an early version of the system, spurious jumps were caused by the user dashing back and forth on the playfield. This was corrected by mapping the measured velocity directly to avatar velocity when the avatar is near ground. Note that although the direct velocity mapping is more robust to noise, the acceleration is still needed because using the velocity mapping when the avatar is in the air causes unnatural mid-air jumps if the user is already moving on the ground after a jump.



Figure 5. The user study setup, shown from the point of view of the webcam used for the computer vision.

Near the ground, there is a non-exaggerated region to further eliminate noise and spurious jumps. The avatar's vertical position is obtained as

$$y = c_y \max(0, y_p - h_0) + \min(y_p, h_0), \quad (2)$$

where  $c_y$  is the vertical motion scaling factor that controls the amount of exaggeration,  $h_0$  is the threshold of the non-exaggerated region and  $y_p$  is the vertical position given by the physics.

Note that the vertical motion of the avatar is not completely realistic even when  $c_y=1$ . The virtual gravity can be chosen arbitrarily and we only have an approximation of the user's real mass center. When computing the mass center of the user pixels in the profile view, your legs and arms weight relatively more than in real life and can thus throw you upwards even while airborne.

#### INFORMAL TESTING: THEATER AND TELEVISION

Kick Ass Kung-Fu has been used in the theater setup in Figure 2, in a television show and at several smaller events. In the television show, Finnish celebrities played the game against cartoon-animated virtual celebrities.

According to our experiences, Kick Ass Kung-Fu works well as entertainment, both for the player and for the audience. In the theater, the stage was open for everyone. The installation attracted audience and people were not timid but queued up to try out the game. In the television show, the game worked well for people with no martial arts background, for example, with football-style kicks. The game has also been popular at parties, even with people in their thirties or forties wearing suits, especially after a few drinks.

Whether people have the courage to go on stage depends a lot on the previous player. We have usually had a host who at times shows some nice moves, but also fools around so that people don't feel too self-conscious. Beginners also feel more confident with some kind of a weapon or playing together with a friend. It has been popular to play the game together with a friend, back to back, each player fending off

the enemies at his or her side with a styrofoam sword that doesn't hurt if you accidentally slash at the other player.

The enemies in the game have a simple artificial intelligence, but they can also be controlled by a person in Wizard of Oz manner using a keyboard. When using the game in a special event like the television show, it is easier to have a human act out the enemies than to implement an artificial intelligence system with a certain character.

#### USER STUDY: ENTERTAINMENT OR EDUCATION?

To find out whether the game has other than entertainment value, we organized a user study with 46 martial arts practitioners. The test setup is shown in Figure 5. One researcher used the computer and directed the test. Another researcher took notes and operated a DVD recorder that recorded the test via a camera positioned so that both the playfield and one screen were visible.

#### Participants

46 martial arts practitioners participated in the study, recruited by sending email to local martial arts clubs. 35 participants were male and 11 female. The ages ranged from 18 to 35 years with a mean of 26. 21 of the participants had a background in taiko, 13 in capoeira, 12 in karate and 8 in acrobatics. In addition to that, there were participants with experience in taekwondo, aikido, ju-jitsu, judo, kickboxing, hapkido, parkour, kungfu, krav maga, mixed martial arts, kali and kendo. The amount of experience in a single sport ranged from 1 to 16 years with a mean of 3.3. Many participants practiced more than one sport.

The participants were asked to sign in as groups of 2-3 persons. Group testing was used to stimulate discussion [27]. In order to find out how the game differs from various martial arts, it was also considered useful to have people who train together analyze each other's performance in the game.

#### Procedure

The test was divided into three parts: introduction, playing the game and an interview.

In the introduction, the participants filled in a background survey, including permissions to use video and photo material taken during the test. The participants were also shown an introductory video of the game to ensure an equal level of prior knowledge. A link to the video was distributed in the recruiting email to promote the test.

After the introduction, the participants took turns in playing, divided into the following phases:

1. The player was told to try moving on the playfield. Horizontal motion scaling values 1, 1.5 and 2 were tried in random order. We wanted to measure whether and how much people wanted their motion to be exaggerated.
2. The player was told to jump and try out techniques involving jumping. Three settings for vertical motion were tried in random order. The first 34 subjects tried an exaggerated but realistically fast jump setting and two slower jump settings, one with smaller gravity and one with slow motion at 0.75 times normal speed. For the last 12 subjects, the setup was changed. They tried a setting where the camera texture's vertical position was constant so that motion was completely realistic, and two exaggerated settings with  $c_y$  values 2.5 and 5. The test instructor could jump two and four times his own height with the exaggerated settings, respectively.
3. Two rounds to death (running out of energy) were fought, with one of the screens showing the same or a horizontally mirrored view compared to the other screen, changed between the rounds. One round took approximately 1-2 minutes. Heart rate was measured after each round. The player was also asked which screen setting felt more natural and why. This was done because in the theater show, players often felt lost for a moment after changing the screen they were facing. The authors themselves tried out both normal and mirrored screens and had differing preferences. The enemy was controlled so that it jumped over the player, making him or her change the attack direction.

After the two rounds, the player was changed. After all participants had played two rounds, everyone had one more round with a flying enemy.

We gave a heart rate monitor to one person in each group and some participants also brought their own ones. The participants with a monitor played last so that an estimate of their resting heart rate could be measured when sitting down and relaxing while the others played. It should be noted that the estimate is not very accurate due to stress caused by the test situation and possible physical exercise preceding the test.

Finally, the participants were interviewed, discussing the following topics in free order:

- Good and bad sides of virtual training in a Kick Ass Kung-Fu style setting.
- User interface – the screens, the playfield, any difficulties or anomalies detected.
- Moving and technique – how the game differs from the martial arts the participants practiced, for example, in terms of use of force, pose and movements, and applicable techniques.
- Future improvements – what should be done to make the game a viable training environment for a longer time.

After the interview, there was an informal playing and experimenting session, if there was time before the next group. Each group was allocated two hours of time.

## RESULTS AND DISCUSSION

In the following, we summarize the results, observations and interview data, discussing various aspects of the game.

### Motion Exaggeration

The distributions for the motion exaggeration parameters are shown in Figures 6-8.

The distributions indicate that the majority of the subjects liked to have their motion exaggerated. Considering the vertical acceleration, the subjects made such remarks as *"Jumping is always fun."*, *"That backflip looked really cool on the screen."* and *"This is how I would like to see myself."* Often, when trying out realistic jump settings after exaggerated ones, the subjects burst out laughing at how feeble they looked on the screen. It was commented that that *"You start to believe in what you see."*

Too much horizontal exaggeration can make estimating distances difficult. In particular, taiko techniques usually consist of a combination of moving and attacking, such as first stepping towards the opponent and spinning around before kicking or punching. With  $c_x=2$ , taidakas often ended moving past the enemy when attacking.

In general, the more athletic the player, the less exaggeration was preferred. Two of the first 34 participants had so much thrust that they practically flew into orbit when performing backflips or other air acrobatics. For this reason, realistic vertical motion was given as an option for the last 12 subjects. Currently, an additional nonlinearity has been added to the physics so that jump height saturates at a reasonable level. This also prevents vertical exaggeration problems when attacking enemies on the ground with jump kicks. If the jump takes much more time than in reality, the kick gets executed too early, after which the avatar just falls on the enemy.

An acrobatics instructor felt that the motion exaggeration *"does the same visually as a springy gymnastics floor does physically"*. The metal springs under such a floor make it easier to launch up high in the air. However, the instructor

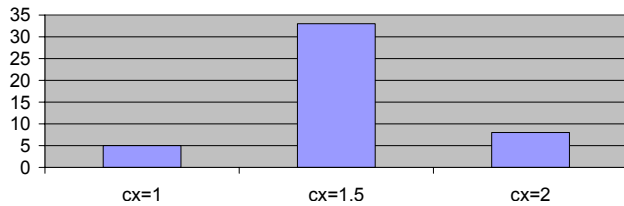


Figure 8. Horizontal motion exaggeration preferences of all subjects.  $c_x=1$  corresponds to realistic motion.

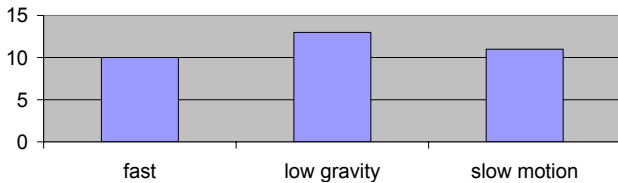


Figure 6. Vertical motion preferences of the first 34 subjects, comparing different exaggerated settings.

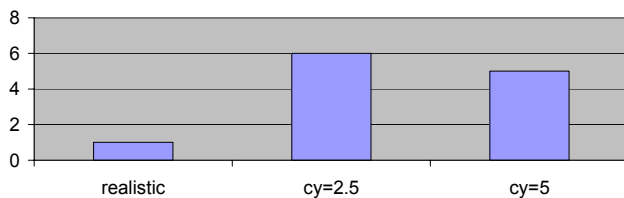


Figure 7. Vertical motion preferences of the last 12 subjects, comparing different jump heights.

preferred our floor made of two centimeters of styrofoam with vinyl coating. He has a black belt in taekwondo and he thought that it would be easy to sprain an ankle when doing a jump kick on too springy a floor.

To summarize our experiences, making motion exaggerated yet natural can require a lot of fine-tuning even when using a minimal subset of game physics. However, the presented control system creates good results even without a real 3D tracking system with multiple cameras. The jump system adds airiness to leaps and easy acrobatic moves, such as a cartwheel where you start from a crouching position and throw your legs upwards. The system encourages you to use your whole body dynamically when jumping, since pulling your legs up after launching gives you extra height. This is important in real-life jump kicks, because pulling your supporting leg up during the kick makes you less vulnerable to a countering leg sweep.

### Slow Motion

Slow motion can be used to achieve a similar visual effect as with low gravity, but with the exception that the timing of jump kicks is easier since the kicks get slowed down in addition to the jump trajectory. 11 out of 34 subjects preferred the slow motion jumps.

We found too much slow motion to be problematic in combat, since people feel that they should wait for the avatar when it lags behind in a jump. However, the heavy slow motion was considered useful in free training. Two groups stayed after the test to practice movie-style jump

kicks at targets placed high in the air. They found it both fun and useful, since the slow motion allowed them to inspect their technique in more detail.

### Heart Rates: Intensive Training

Heart rate data was gathered from a total of 20 subjects. Table 1 summarizes the heart rates in rest and after fighting one round with the ground based and flying enemies. The last column of the table shows the median of heart rates as a percentage of maximum heart rate  $HR_{max}$ , estimated as 220 minus each person's age [22]. This percentage is often used to express intensity level of physical training. The results show that the game can be regarded as intensive physical training, as intensities above 70% are recommended for fitness training for people who exercise regularly [22].

	Median (BPM)	Max (BPM)	Min (BPM)	Median % HRMax
Resting	72	87	43	37%
Enemy on the ground	171	191	157	90%
Enemy flying	179	200	165	93%

Table 1. Summary of the heart rates measured in rest and after one round with a ground-based and flying enemy.

Fighting a flying enemy requires more jumping, which can increase the intensity level. The highest heart rate of the study, 200 BPM, was also measured with the flying enemy. Statistically, there appears to be no significant difference in the heart rates against the different enemies.

In the interviews, 19 subjects considered the game more exhausting than combat in the martial arts they practiced. The reasons mentioned were that you get more enthusiastic and carried away, the virtual enemies do not get tired and in real combat there is more lurking and waiting for openings.

### Screens and the Profile View

The preferences regarding mirroring one of the screens horizontally were divided quite equally. The effect of the mirroring was not always clear and 9 users out of 46 did not have an opinion. Of the 37 users who had an opinion, 21 preferred the mirrored setting. However, the opinions show no statistical significance, since they could have been obtained randomly with a probability  $p=0.09$  from a binomial distribution  $\text{Bin}(37,0.5)$ .

Although statistics gives no results, we find the users' comments on the mirroring interesting. Most of the subjects felt that they could adapt to both configurations, but some had strong opinions. It was found that some people think of directions in terms of screen coordinates, along the lines of "The enemy is on my right on the screen. Thus, I must move forward." This kind of right-to-forward mapping is consistent between the screens only if one screen is mirrored. On the other hand, some people think in real-world coordinates, such as "I thought that since the enemy was in the direction of that wall (points at a wall in the test

space), *I must move forward*". This kind of mapping between real and virtual directions is consistent only if both screens are the same.

Despite the differences in perceiving directions, many users commented that the profile view helps in spotting errors in pose and technique. Only one out of the 46 users got momentarily confused so that she started attacking in screen coordinates, kicking towards the camera instead of in the direction of the playfield.

The subjects often felt confused when switching their gaze from one screen to another, after which a moment of reorientation was needed, especially in the beginning. It was commented that if your opponent jumps over you in real life, you can follow him or her with your eyes. If you don't follow, you lose him or her from sight, which signals that you should turn around. In the game, you can still see the enemies on the screen even if they are behind you.

Many participants felt that the screens were too high, so that it was difficult to keep your eyes at your opponent when doing cartwheels or other moves where it is difficult to look up. As a capoeirista remarked, you see at least your opponent's feet when doing a cartwheel in real life. It appears that the screens should be placed as close to the floor as possible and the virtual camera should be pointed so that the ground level is near the bottom of the screen.

#### Observations on Moving and Technique

In the interviews, the most often repeated observation related to moving and technique was that you tend to move carelessly in the game, which has its good and bad sides. For example, the virtual opponent cannot grab a hold of your foot so that it is easy to forget to return your foot properly after a kick. A karateka noted that he lunged forward more than usual when punching, which would make him open for counterattacks in real life.

On the other hand, many users remarked that it is good that the techniques are quick and relaxed, and that you get to do a lot of repetitions. Basically, you don't have to worry about hurting yourself or the opponent, which allows you to try out acrobatic and hard-to-control techniques. However, you should be careful not to overextend your knees or other joints, which applies to all training without a physical target.

The biggest drawback of the game seems to be that you can't dodge to the sides or move otherwise three dimensionally, which is integral to some martial arts, such as taiko. For people inexperienced in martial arts, the usual reflex is to dodge by stepping back, but dodging to the side or front is much more useful if you want to get to a good position for a counterattack. Dodging is also made difficult by the delay caused by camera drivers and display double buffering. One user remarked that the delay is visible in sharp punches. The delay also appeared to cause many users crouch down too late when trying to dodge bullets fired by the flying enemy.

The 2D playing field seems to be more suitable for less three-dimensional martial arts, such as taekwondo and kendo. In kendo, a form of Japanese sword training, you often move on a line and you only rarely slash sideways with the sword. In taekwondo, you dodge to the sides, but you also practice a lot of combinations of kicks on a line.

An obvious difference between the game and real martial arts is the lack of physical contact. You cannot grab or throw the opponent. It was suggested that computer vision could be augmented with wearable grabbing sensors, which could be implemented quite easily, but the real challenge is to provide realistic force feedback so that you cannot just grab your virtual 200 pound opponent and shake him or her in the air like a rag doll.

#### Model Based vs. Low Level Computer Vision

In addition to dodging and grabbing, the game also currently lacks the possibility of blocking attacks, for which a more sophisticated computer vision system discerning body parts would be needed. In the interviews, it was noted that the lack of blocking makes you forget to keep your hands in a guarding position.

There are two basic approaches to designing computer vision for a game like Kick Ass Kung-Fu. You can control the game with the parameters of a model that is fitted to the image. A variety of models have been proposed, such as skeletons [21], deformable contour templates [6], color blobs [29], and probabilistic models [24]. Instead of a model, you can also interpret low-level image features using application-specific prior knowledge and assumptions [11][14].

In our case, the low-level features are the location of the mass center and optical flow. We abandoned the model-based approach in order to allow the player use any technique with no limitations, including a variety of weapons. In model-based vision, the model would have to contain information of the weapons as well as the user. Generic low-level vision also simplifies the hardware setup. For correctly fitting a several degree of freedom 3D skeleton to a player that kicks, crouches, spins and crawls on the floor, you need considerably more processing power than for the simple optical flow and probably also additional hardware, such as wearable markers and multiple cameras.

In the theater setting, people spontaneously used weapons, such as chairs and scarves. The optical flow registered the motion of the weapons well, although the weight and thus the impact of the weapons could not be analyzed realistically. A fast-moving scarf was therefore a fun and effective weapon. In the user study, the game was successfully played with a kendo practice sword and variety of wooden sticks. Perhaps the most fun weapon was a 10cm wide mini-frisbee. One user kept a pile of such frisbees in his hand and threw them at the opponents like throwing-



stars. The frisbees were big enough for the camera to detect them and light enough not to damage the screens.

### CONCLUSIONS AND FUTURE WORK

We have presented the design of Kick Ass Kung-Fu, an artificial reality martial arts game installation, together with the results and observations from testing the game. In its current state, the game seems to work well for entertainment and fitness applications, based on the reactions of the users and the high heart rates measured. The dual screen profile was found to work and only one of the 46 user study participants had problems with the 90 degree rotation between world and screen coordinates. The motion exaggeration and dynamic slow-motion methods were found to add excitement to the game. The methods could also be applied to other game genres, such as platform jumping games.

Kick Ass Kung-Fu can be seen as part of an emerging phenomenon of *computer games as performances*. Previous examples include dancing games, such as Dance Dance Revolution [1], and more recently, games played by singing [15]. The common factor is that playing a game can actually be great entertainment not only for the player, but also for people around, which can have interesting social implications. If playing a game requires and develops athletic or musical skills, skilled players can obtain an idol status similar to pop stars or athletes, appealing to a much larger audience than, say, first person shooter gurus. In the future, we will try to promote the forming of user communities, for example, by having videos of best players automatically uploaded to the Internet.

According to our experience, Kick Ass Kung-Fu can augment and motivate martial arts and acrobatics training, at least at beginner and intermediate levels. For more advanced technique and combat training, the users wanted the game to be more realistic. Our low-level computer vision system allows multiple players and practically all weapons, but in the future, we will consider model based computer vision for increased realism. Using a skeleton model would enable you to block attacks and interact more realistically with the virtual environment. However, realism poses challenges in interaction design as well as technology, since it is nontrivial to control an avatar that reacts to the environment. Currently, the mapping between the user and the avatar is one-to-one, except with respect to vertical position. The mapping breaks if the avatar can trip over virtual obstacles without real-world counterparts, or if impact forces are applied to the avatar.

Although motion exaggeration was found fun in the game, many experienced martial artists requested a training mode with realistic motion and enemies. Related to this, many participants suggested a two-player mode as a fun and motivating addition to the game. Each player should have his or her own playfield to avoid real contact. Playing against another human could also make the game

interesting for a longer time, since your opponents develop and learn with you.

In the interviews, the game setup was considered promising for practicing predetermined sequences of techniques (kata in karate). Kata training often focuses on fluidity and rhythm, for which it could be useful to have a group of enemies appear in a predetermined order, possibly in rhythm with music.

A group of experienced taido practitioners suggested that it would be useful to train your reactions and timing against a pre-recorded video opponent. For example, in dodging and counter-attacks, it is important not to act too late or early. The video playback speed could be adjusted so that you start practicing at a slow tempo and then increase speed as you start to master the moves. However, camera and display delays can be a problem for this kind of training.

Many users complained that the game encourages you to use only an effective and fast subset of techniques. The same applies to real combat, but for example taido competition combat rules prohibit straight kicks and punches without certain preceding motion patterns, which makes the combat more jazzy and dynamic. Creative playing could also be encouraged by a style based a scoring and rewarding system. Such systems exist in traditional video games, such as the Lord of the Rings Two Towers game, where successful combinations of attacks produce higher scores. For a game with a video avatar, it could be natural to have model choreographies performed by an expert, showing an optimal way of demolishing a horde of attackers.

The Kick Ass Kung-Fu user interface can be considered to develop motor skills and fitness, but from an educational point of view, this is only part of the experience. Martial arts are about both mind and body, about harmony and respecting your fellow trainees. In the future, the philosophical aspects should be incorporated into the game, for example, in the form of narrative.

Information and audiovisual material about Kick Ass Kung-Fu can be found at <http://kickasskungfu.net>.

### ACKNOWLEDGMENTS

This work has been supported by HeCSE graduate school, Nokia Foundation, Tekniikan Edistämissäätiö and KAUTE. The authors would like to thank Sonja Kangas, Tapio Takala, the anonymous reviewers, and all who have played the game and helped us with their comments. Most importantly, we are grateful to our martial arts teachers Peter Nysten, Kurt Berger, Markus Vainio-Mattila, Hasse Sjöholm, and Jon Sundwall.

### REFERENCES

1. Konami Annual Report 1999, <http://www.konami.co.jp/en/ir/financialinfo/annualreport/report1999/am01.html>, link checked 4<sup>th</sup> Jan 2005

2. Sony Eye-Toy, <http://www.eyetoy.com>, link checked 9<sup>th</sup> Sep 2004
3. OpenCV, <http://sourceforge.net/projects/opencvlibrary/>, link checked 9<sup>th</sup> Sep 2004
4. Ajdari, J., Aulaskari, I., Brown, J., Edutainment in learning environments, <http://www.helsinki.fi/~jbrown/tao/rap7.html>, link checked 4<sup>th</sup> Jan 2005
5. Baecker, R., Miller, D., and Reeves, W. Towards a Laboratory Instrument for Motion Analysis. *Computer Graphics*, vol. 15, No. 3, August 1981, pp. 191-197
6. Blake, A., Isard, M., Active Contours The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion, Springer-Verlag London Limited (1998)
7. Chi, E.H., Song, J., Corbin, G., "Killer App" of wearable computing: wireless force sensing body protectors for martial arts, *Proc. UIST 2004*, ACM Press (2004), pp. 277-285
8. Philo Tan Chua, Crivella, R., Daly, B., Ning Hu, Schaaf, R., Ventura, D., Camill, T., Hodgins, J., Pausch, R. Training for Physical Tasks in Virtual Environments: Tai Chi, *Proc. IEEE Virtual Reality 2003*, pp. 87-94
9. Crowley, J.L., Coutaz, J., Bérard, F., Things That See, *Communications of the ACM*, Vol. 43, No. 3, March 2000
10. Cruz-Neira, C., Sandin, D.J., DeFanti, T.A. Surround-screen projection-based virtual reality: the design and implementation of the CAVE, *Proc. SIGGRAPH'93*, ACM Press (1993), pp. 135-142
11. Freeman, W.T., Beardsley, P.A., Kage, H., Tanake, K., Kyuma, K., Weissman, C.D., Computer Vision for Computer Interaction, *SIGGRAPH Computer Graphics Newsletter*, Vol. 33, No. 4, November 1999, ACM SIGGRAPH.
12. Ishii, H., Wisneski, C., Orbanes, J., Chun, B., Paradiso, J., PingPongPlus: design of an athletic-tangible interface for computer-supported cooperative play, *Proc. CHI'99*, ACM Press (1999), pp. 394-401
13. Hämäläinen, P., Interactive Video Mirrors for Sports training. *Proc. nordiCHI 2004*, ACM Press (2004), pp. 199-202
14. Hämäläinen, P., and Höysniemi, J. A Computer Vision and Hearing Based User Interface for a Computer Game for Children. *Proc. 7th ERCIM Workshop "User Interfaces For All"* (2002)
15. Hämäläinen, P., Mäki-Patola, T. Pulkki, V. Airas, M. Musical Computer Games Played by Singing. *Proc. 7th International Conference on Digital Audio Effects (DAFx'04)*, 2004, pp. 367-371
16. Iredale, F., Farrington, T., and Jaques, M., Global, fine and hidden sports data: applications of 3-D vision analysis and a specialised data glove for an athlete biomechanical analysis system, *Proc. Mechatronics and Machine Vision in Practice* (1997), pp. 260-264
17. Krueger, M., Gionfriddo, T., Hinrichsen, K. VIDEOPLACE - An Artificial Reality, *Proc. CHI'85*, ACM Press (1985), pp. 35-40
18. Mueller, F., Agamanolis, S., Picard, R., Exertion interfaces: sports over a distance for social bonding and fun, *Proc. CHI'03*, ACM Press (2003), pp. 561-568
19. Oblinger, D.G. The Next Generation of Educational Engagement, *Journal of Interactive Media in Education*, 2004 (8)
20. Perš, J., Bon, M., Kovačič, S., Šibila, M., Delman, B. Observation and Analysis of Large-scale Human Motion, *Human Movement Science*, 21(2), 295-311, July 2002
21. Plänkers, R. and Fua, P. Tracking and Modeling People in Video Sequences. *Computer Vision and Image Understanding*, 81(3), March 2001
22. Polar A3 heart rate monitor user's manual, Polar Electro (2004)
23. Press, W.H., Teukolsky, S.A., Vetterling, V.T., Flannery, B.P. Numerical Recipes in C++, 2<sup>nd</sup> ed., Cambridge University Press (2002)
24. Sidenbladh, H., Black, M.J., Sigal, L. Implicit Probabilistic Models of Human Motion for Synthesis and Tracking, *Proc. 7th European Conference on Computer Vision - Part I*, Springer-Verlag (2002)
25. Toyama, K., Krumm, J., Brumitt, B., and Meyers, B. Wallflower: Principles and Practice of Background Maintenance, *Proc. International Conference on Computer Vision* (1999)
26. Yamamoto, M., Kondo, T., Yamagiwa, T., Yamanaka, K. Skill recognition, *Proc. Automatic Face and Gesture Recognition* (1998), pp. 604 - 609
27. Wilson, C., & Blostein, J. (1998) Pros and cons of co-participation in usability studies. *Usability Interface*, 4 (4), April 1998.
28. Wren, C.R., Spacarino, F., Azarbayejani, A., Darrel, T., Davis, J., Starner, Kotani, A., Chao, C., Hlavac, M., Russel, K., Bobick, A., Pentland, A., Perceptive Spaces for Performance and Entertainment (Revised), *Applied Artificial Intelligence*, Vol. 11, No. 4, June 1997
29. Wren, C.R., Azarbayejani, A., Darrel, Pentland, A. Pfunder: Real-Time Tracking of the Human Body, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July 1997, Vol. 19, No. 7, pp. 780-785