# VIDEO LOOPING OF HUMAN CYCLIC MOTION

A Thesis

by

HYE MEE CHOI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2004

Major Subject: Visualization Sciences

# VIDEO LOOPING OF HUMAN CYCLIC MOTION

A Thesis

by

HYE MEE CHOI

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Ergun Akleman
(Chair of Committee)

Jianer Chan
(Member)

Carol LaFayette
(Member)

Phillip Tabb
(Head of Department)

May 2004

Major Subject: Visualization Sciences

# ABSTRACT

Video Looping of Human Cyclic Motion. (May 2004)
Hye Mee Choi, B.S., Korean Sahmyook University;
M.F.A., Hongik University
Chair of Advisory Committee: Dr. Ergun Akleman

In this thesis, a system called *Video Looping* is developed to analyze human cyclic motions. Video Looping allows users to extract human cyclic motion from a given video sequence. This system analyzes similarities from a large amount of live footage to find the point of smooth transition. The final cyclic loop is created using only a few output images. Video Looping is useful not only to learn and understand human movements, but also to apply the cyclic loop to various artistic applications. To provide practical animation references, the output images are presented as photo plate sequences to visualize human cyclic motion similar to Eadweard Muybridge's image sequences. The final output images can be used to create experimental projects such as composited multiple video loops or small size of web animations. Furthermore, they can be imported into animation packages, and animators can create keyframe animations by tracing them in 3D software.

To my family

# ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Ergun Akleman, for sharing his knowledge and guidance, which helped me accomplish this thesis and led me through the difficulties of this research. I would also like to thank my committee member, Carol LaFayette, who gave me good critiques and fresh ideas and encouraged me to finish the animation *March of the Chicken-Man*. Special thanks to my committee member, Dr. Jianer Chen for providing his help throughout this research.

A very special thanks to all the people who performed the shooting video to input video sequences in my thesis program. My sincere gratitude goes to all the faculty and staff of the Viz Lab, who have provided advice and tremendous help. I would also like to thank my friends at the lab, who helped me to overcome all difficulties of the past three years and believed in my potential and talents.

Finally, I would like to thank my family for their continuous support to achieve my goal and dream and for their unconditional love which has helped me all through my life.

# TABLE OF CONTENTS

CHAPTER Page

I INTRODUCTION . . . . . 1

I.1. Motivation . . . . . 1
I.2. History of the Study of Motion . . . . . 3
I.3. Animation Methods . . . . . 11

II RELATED WORK . . . . . 15

II.1. Video Speech Animation . . . . . 15
II.2. Video Motion Capture . . . . . 16
II.3. Video Texture . . . . . 16
II.4. Video Sprite Animation . . . . . 18
II.5. Cartoon Capture and Retargeting . . . . . 20

III METHODOLOGY . . . . . 21

III.1. Shooting Video . . . . . 21
III.2. Bounding the Box . . . . . 23
III.3. Cropping Frames . . . . . 23
III.4. Finding Identical Frames . . . . . 26
III.5. Creating Loop . . . . . 27

IV IMPLEMENTATION AND RESULTS . . . . . 29

IV.1. Video Looping Tool . . . . . 29
IV.2. Digital Cyclic Motion Library . . . . . 30
  IV.2.1. Walk . . . . . 34
  IV.2.2. Run . . . . . 37
  IV.2.3. Hop and Jump . . . . . 39
  IV.2.4. Carrying a Heavy Object . . . . . 43
IV.3. Compositing Multiple Video Loops . . . . . 45
IV.4. Rotoscoping in 3D Animation . . . . . 48

V CONCLUSION AND FUTURE WORK . . . . . 49

REFERENCES . . . . . 51

APPENDIX 1 . . . . . 54

VITA . . . . . 57

# LIST OF FIGURES

FIGURE Page

1 Eadweard Muybridge, *Horse in Motion.* [10] . . . . . . . . . . . . . . . . 3

2 Eadweard Muybridge, *Muybrige's Human and Animal Locomotion.* Plate 46: Child running. [19] . . . . . . . . . . . . . . . . . . . . . . . 4

3 Eadweard Muybridge, *Animals in Motion.* Plate 115: Dog walk. [18] . . . 4

4 Marey's fusil photographique. [1] . . . . . . . . . . . . . . . . . . . . . . 6

5 Marey's black costume for geometric chronophotography, 1884. [1] . . . 6

6 Etienne-Jules Marey, (a) Joinville soldier walking, 1883. (b) Joinville soldier running, 1883. (c) Joinville soldier running, 1883. [1] . . . . . . . 7

7 Etienne-Jules Marey, (a) Mounting, 1891. (b) Dismounting, 1891. (c) Pushing a cart, 1891. [1] . . . . . . . . . . . . . . . . . . . . . . . . . . 8

8 Eadweard Muybridge, *Ascending and Descending Stairs*, from the series, Animal Locomotion (1884-85). [19] . . . . . . . . . . . . . . . . . . 9

9 Marcel Duchamp, *Nude Descending a Staircase, No. 2, 1912.* [1] . . . . . 9

10 Etienne-Jules Marey, *Flight of Gull*, 1886. [1] . . . . . . . . . . . . . . . 10

11 Giacomo Balla, *Swift: Paths of Movement + Dynamic Sequences*, 1913. [1] . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

12 Use of live action in *Snow White.* Reprinted with permission from [9] . . 12

13 Generated video texture of candle flames. Reprinted with permission from [24] . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

14 The final example of a complete fish tank. Reprinted with permission from [24] . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18

15 Overview of video sprite animation pipeline. Reprinted with permission from [25] . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19

FIGURE Page

16 The generated hamster and fly animations. Reprinted with permission from [26] . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19

17 System overview diagram. . . . . . . . . . . . . . . . . . . . . . . . . 22

18 Compute alpha and RGB values using Video Looping system. (a) Computing alpha values and the bounding box size of a character. (b) Calculating color values. . . . . . . . . . . . . . . . . . . . . . . . . . 24

19 Cropping final output images based on the bounding box. . . . . . . . . . 25

20 Image similarities and transitions. $D_{ij}$ is the distance function. Reprinted with permission from [25] . . . . . . . . . . . . . . . . . . . . . . . . 27

21 Finding good transitions in a pendulum sequence. . . . . . . . . . . . . . 27

22 Captured screen shots of Video Looping program in Irix. (a) Probability matrix. (b) Play of a final cyclic loop using OpenGL. . . . . . . . . 28

23 Captured screen images of Digital Cyclic Motion Library. . . . . . . . . 32

24 Combining two different final loops by skipping some frames. (a) Select frame 1-20 from the final loop of Walk 1. (b) Select frame 18-31 from the final loop of Walk 2. . . . . . . . . . . . . . . . . . . . . . . 33

25 Dog walk. Total frames: 36, best frames: 14-6, range: 8. . . . . . . . . . 34

26 Child walk. Total frames: 206, best frames: 139-119, range: 20. . . . . . 35

27 Female walk. Total frames: 131, best frames: 88-55, range: 33. . . . . . . 35

28 Male walk. Total frames: 70, best frames: 43-9, range: 34. . . . . . . . . 36

29 Child run. Total frames: 77, best frames: 24-9, range: 15. . . . . . . . . . 37

30 Female run. Total frames: 161, best frames: 109-80, range: 29. . . . . . . 38

31 Male run. Total frames: 131, best frames: 90-67, range: 23. . . . . . . . . 38

32 Male hop. Total frames: 145, best frames: 79-61, range: 18. . . . . . . . 39

33 Female hop. Total frames: 79, best frames: 34-18, range: 16. . . . . . . . 40

FIGURE Page

34 Female hop with one leg. Total frames: 113, best frames: 58-42, range: 16. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40

35 Female jump. Total frames: 296, best frames: 102-66, range: 36. . . . . . 41

36 Male jump. Total frames: 206, best frames: 119-74, range: 35. . . . . . . 42

37 Female carrying an object. Total frames: 143, best frames: 59-32, range: 27. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43

38 Male carrying an object. Total frames: 112, best frames: 50-14, range: 36. 44

39 Shooting and bluescreening technique. (a) Input frames with the bluescreen: 0-230. (b) Final skipped frames: range 14. (c) Extracted the character from the bluescreen. . . . . . . . . . . . . . . . . . . . . . . . 46

40 The process of creating 2d animation, ’March of the Chicken-Man’. (a) Duplicating a loop of character’s walk cycle. (b) Digital illustration for background. (c) Final composited new shot. . . . . . . . . . . . 47

41 Importing the final images to 3D software, Maya®. . . . . . . . . . . . 48

42 Determination of the bounding box when the background is not the bluescreen. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55

43 Extracting a character from the background. . . . . . . . . . . . . . . . . . 56

# CHAPTER I

# INTRODUCTION

## I.1. Motivation

Currently, computer animations are created using several animation methods. One of the common methods, keyframing is the traditional hand animation technique, which requires the animator to specify key positions for animated subjects. Experienced animators can create subtle details of a character's motion using a complicated process of setting up keyframes and adjusting them. However, creating high quality keyframe animation is labor intensive and requires plenty of time and a high level of expertise.

Recent advances in motion capture, performance animation and the incorporation of scripting into keyframe systems are beginning to break the ascendancy of keyframing and open up other ways of creating animation. Motion capture is a process for capturing motion from a performer [4, 21]. It employs special sensors called trackers to record the live motion of an actor. In the process, the system records accurate details and variations of an actor's movement. However, it requires a fast and easy setup, quick calibration, and quick turnaround from the time of motion capture to the finished data set.

Even though many techniques have been developed to edit motion capture data [15], the information captured is difficult to edit and requires substitution of massive amounts of data with usable data. The other aspect of motion capture data is that the same motion cannot be easily executed twice [12, 21]. Once live motion is recorded, the same cycle of captured data cannot be repetitively generated [20]. Due to these reasons, motion capture has been mainly used in large computer graphic (CG) production studios that can afford

---

The journal model is *IEEE Transactions on Visualization and Computer Graphics.*

expensive equipment. For individual animators and small companies, there is a need for inexpensive digital image acquisition hardware and cheaper, more effective techniques.

More and more production studios are combining techniques from keyframe animation systems, plug-ins, custom code, and motion capture in order to improve the quality and efficiency of their animation. However, human motion study is still challenging because all the degrees of the various configuration of the legs, arms, torso, and head are different and subtle. In accordance with recent trends and research, efficient techniques of motion analysis for computer animation are necessary to create realistic characters through learning about and understanding human movement. For instance, a video-based technique is one of the potential approaches to creating realistic animation. The objective of one video-based technique is to generate new image sequences through rearranging original video footage [8]. It is a practical method to solve the problems of excessive labor, limited time, and limited budget.

The goal of this thesis is to study human cyclic motions using advances in CG technology and to create a simple video-based technique for computer animation. This technique does not require any expensive equipment or special expertise. It allows to easily collect and manipulate human cyclic motion data from original video. It also provides a simple way to generate a cyclic loop from original video sequences and preserves the motion features of original video by generating a loop. More specifically, by analyzing similarities from a huge amount of input video sequences, only a few frames are used to make a loop. Furthermore, this technique enables the generation of a natural cyclic loop based on smooth transitions without any discontinuity.

The final output images of each cyclic motion are demonstrated as photo plate sequences inspired by Eadweard Muybridge's and Etienne-Jules Marey's sequential method. These consecutive sequences will be a good reference for animators who want to create elaborate motions.

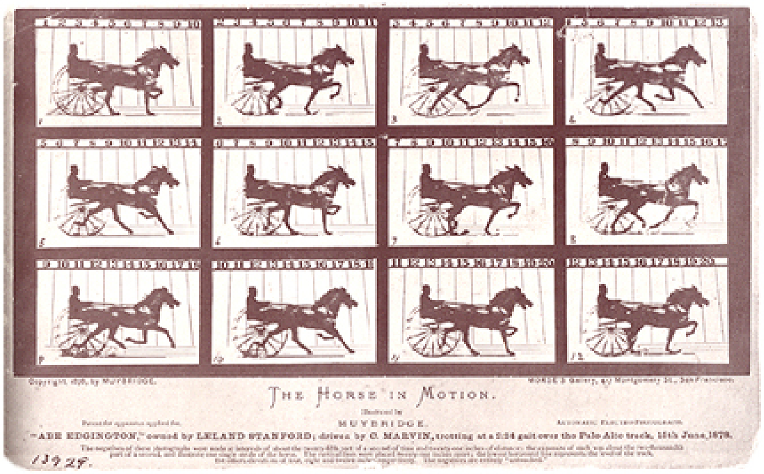


Fig. 1. Eadweard Muybridge, *Horse in Motion.* [10]

## I.2. History of the Study of Motion

During the mid 1900s, the invention of photography provided artists new ways to visualize still and moving images. In 1878, one of the pioneers of *chronophotography*, Eadweard Muybridge, successfully took the first serial photographs of fast horse motion [1, 10]. Muybridge set up twelve cameras, equipped with fast stereo lenses and an electrically controlled mechanism to operate the camera's special shutters. His first publication called "The Horse in Motion," included twelve pictures of a horse, that were taken in about half a second (see Figure 1). Undeniably, this was the result of his long work and intense effort in designing and assembling the equipment used. Thus, Muybridge continued this sensational photographic experiment by expanding the setup to twenty-four cameras. He was able to obtain successful results of movement of more animals and athletes from the Olympic Club of San Francisco. As shown in Figure 2 and Figure 3, he showed continuous rapid movements of men and animals by placing the resulting images next to each other [17, 18, 19].

His eleven-volume studies entitled “Animal Locomotion”, published in 1888, presented the idea of representing motion in time to many artists [1, 10, 18].

Fig. 2. Eadweard Muybridge, *Muybrige’s Human and Animal Locomotion.* Plate 46: Child running. [19]

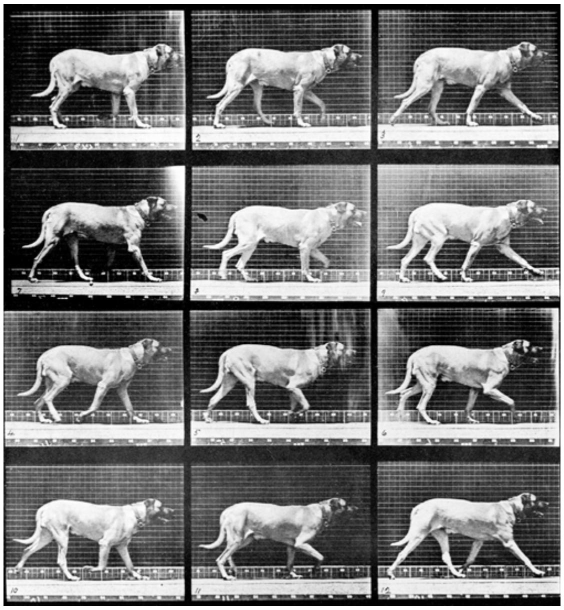


Fig. 3. Eadweard Muybridge, *Animals in Motion.* Plate 115: Dog walk. [18]

Muybridge's work later influenced the physician and physiologist Etienne-Jules Marey who recorded the movements of for example, a bird in free flight, seahorses and human beings [1, 7]. To improve his studies, Marey developed precise recording instruments, such as the *photographic gun*, which recorded twelve exposures on one plate (see Figure 4). For the human movement, the subjects wore black costumes with metal strips or white lines to help capture motions as they passed in front of black backdrops (see Figure 5). As shown in Figure 6 and Figure 7, he created a synthesis of movement in time and space, rather than an analysis in sequential rows of images as Muybridge had done. His work had an impact in many technical and scientific areas, such as cardiology, microscopic cinematography, biomechanics, physical education, and aviation.

While Muybridge continued his motion studies, the influence of his photographs was felt everywhere, especially upon the arts. His studies revealed the subtle details of the human and animal motion that were hard to catch with the naked eye. Many futurists such as Giacomo Balla and Umberto Boccioni began to focus attention on experiments which expressed movement and human gestures. Marcel Duchamp's painting, *Nude Descending a Staircase, No. 2* is one of the greatest example. Duchamp abstracted human motion in his own explorations of time [14, 23] - the fourth dimension - with inspiration from Muybridge's motion study (see Figure 8 and Figure 9). Balla also incorporated visual discovery of Marey's chronophotographs, *Flight of Gull* (see Figure 10 and Figure 11) in his work [1].

Although Marey had no interest in constructing an illusion of movement, his single-camera system paved a way to cinematography [7]. Consequently, his desire to translate the movements of humans and animals into a visible language contributed to the replication of movement in the making of films.

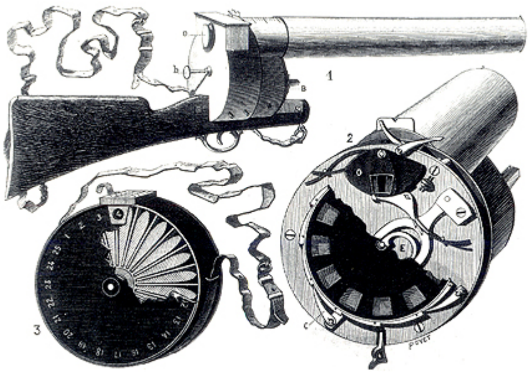
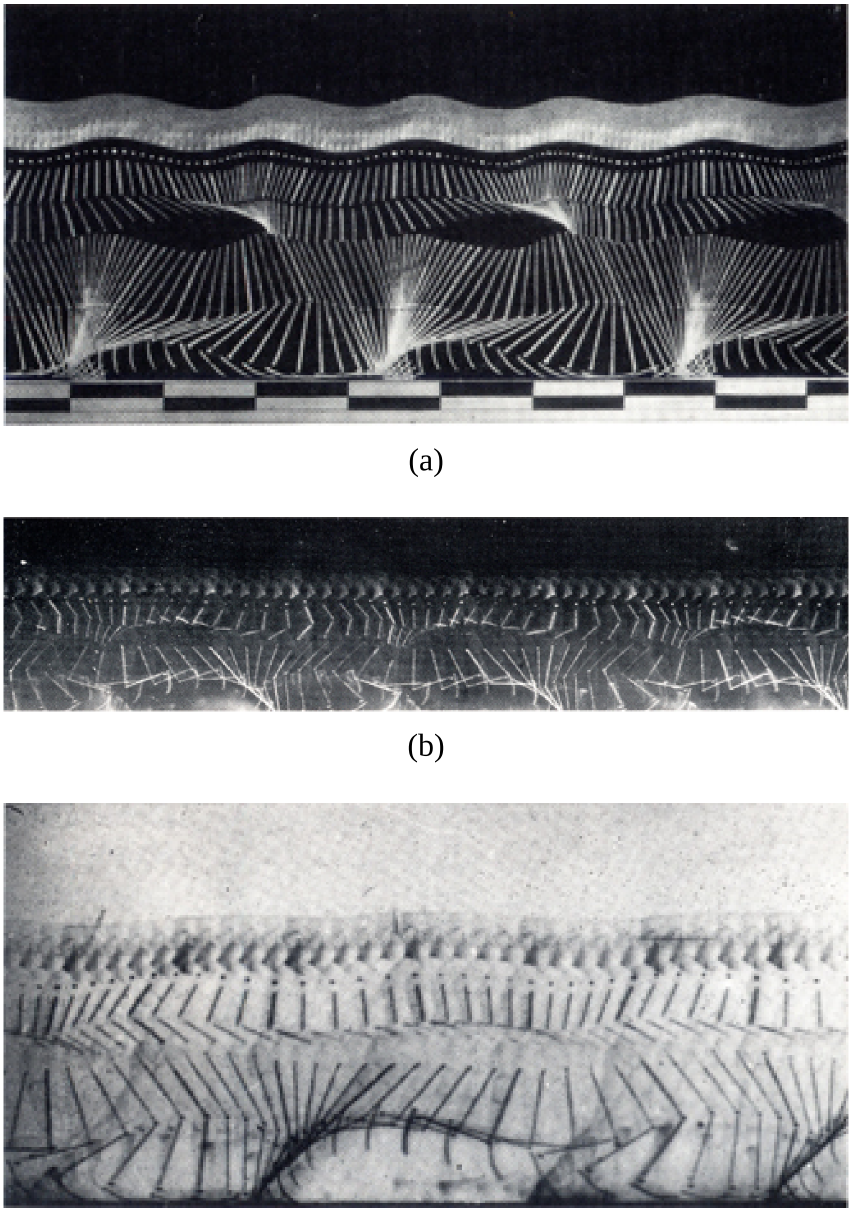


Fig. 4. Marey’s fusil photographique. [1]

Fig. 5. Marey’s black costume for geometric chronophotography, 1884. [1]



(a)

(b)

(c)

Fig. 6. Etienne-Jules Marey, (a) Joinville soldier walking, 1883. (b) Joinville soldier running, 1883. (c) Joinville soldier running, 1883. [1]

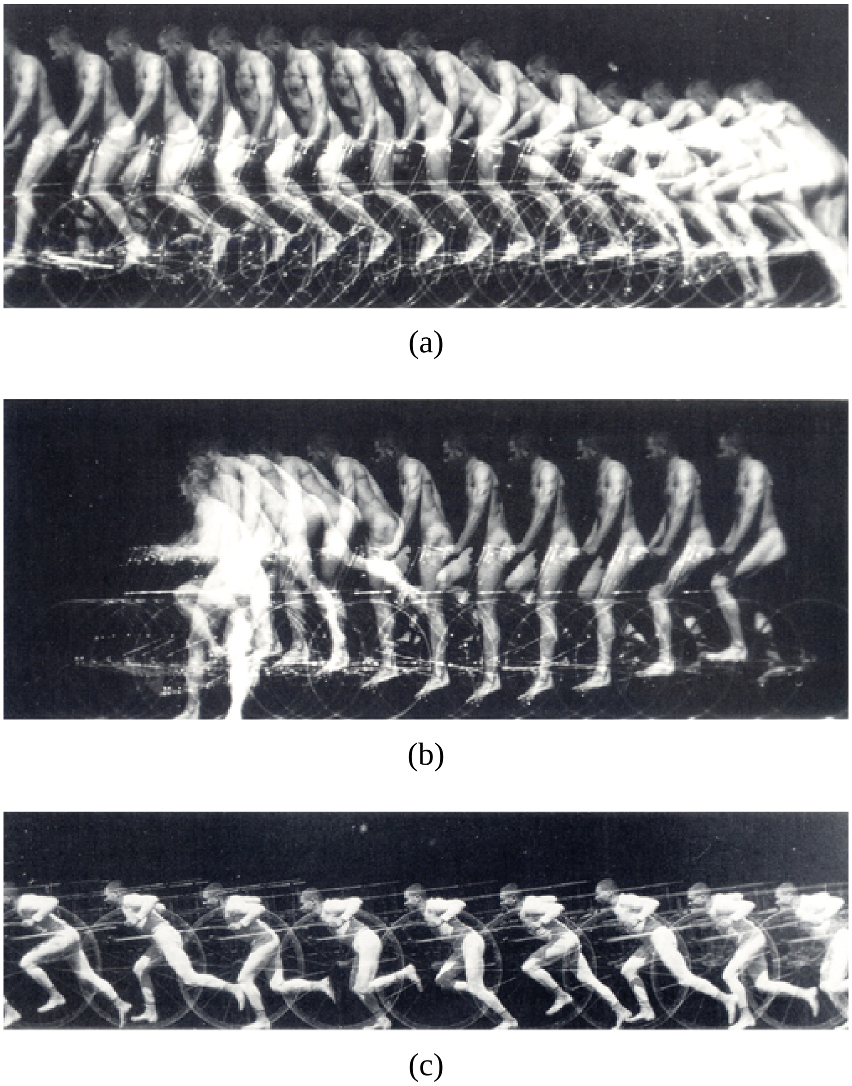


Fig. 7. Etienne-Jules Marey, (a) Mounting, 1891. (b) Dismounting, 1891. (c) Pushing a cart, 1891. [1]

Fig. 8. Eadweard Muybridge, *Ascending and Descending Stairs*, from the series, Animal Locomotion (1884-85). [19]

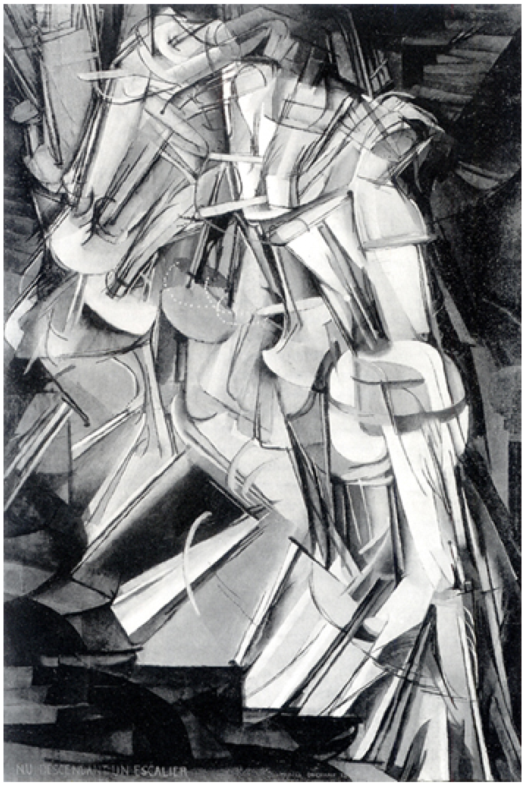


Fig. 9. Marcel Duchamp, *Nude Descending a Staircase, No. 2, 1912.* [1]

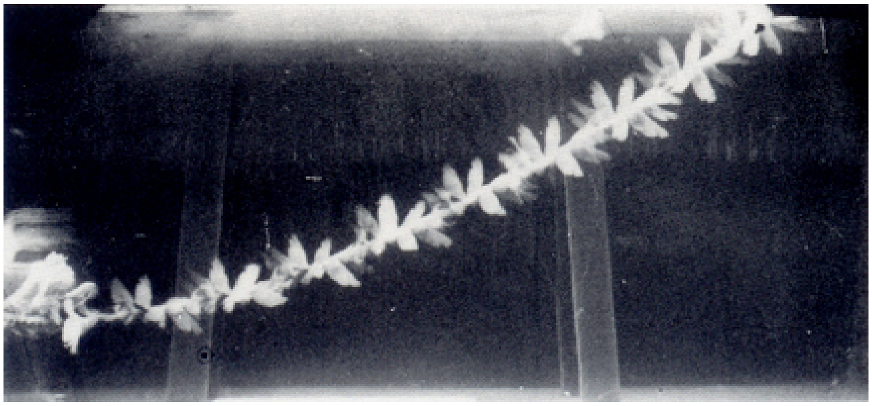


Fig. 10. Etienne-Jules Marey, *Flight of Gull*, 1886. [1]

Fig. 11. Giacomo Balla, *Swift: Paths of Movement + Dynamic Sequences*, 1913. [1]

## I.3. Animation Methods

Muybridge's photographic images and Marey's fundamental human locomotion studies have been a treasure to many artists and film makers who wanted to visualize moving images. Their studies also had a strong impact on animation since precise analysis of human and animal motion helped to create better animations.

Animators have make efforts to study subtle details of movements. In the early twentieth century, live-action filming was one of the solutions. In 1915, Max Fleischer, who owned the studio that brought *Betty Boop* invented a new technique, "rotoscoping." [16, 27] Rotoscoping is the process in which images are traced frame-by-frame onto animation paper [16, 28, 30]. This methodology was used to create "key" drawings, which determine timing and movement. For instance, Walt Disney's studio filmed actors performing movements to create key drawings. As shown in Figure 12, the studio used the resulting footage as a guide for timing, weight and movement to animate characters [16, 28, 29]. It brought success to creating character animations of *Snow White*. Later recording live-action became a part of the animation process [30].

Experienced traditional animators used live-action filming as reference footage for the movements of a character in order to create a character's personality. They determined which details in the footage to exaggerate and enhance. Animators specify how many frames should occur between two poses and then manually set keyframes [13, 21]. This method of "keyframing" is still been commonly used to create animation.

The process of keyframing has been computerized after the digital revolution in the computer graphics industry. The film is digitally scanned onto a disk, and the artist tracks it with high-speed hardware. Animators can then watch and track each frame of the referenced film and draw freehand. They create subtle details of a character's motion using a complicated process of setting up keyframes and adjusting them.

Fig. 12. Use of live action in *Snow White.* Reprinted with permission from [9]

To save time and work and to create subtle details of animation, more complicated and computerized techniques, such as motion capture, were developed at the end of the 1980's. Motion capture is the process of recording live motion and translating it into applicable data by tracking key points on the subject and converting them into mathematical representations. Simply put, it is the process of digitizing a live performance. The process of motion capture can generally be described in three steps:

1. Planning: A list of planned motions is scheduled for capture using storyboards and a shot list.
2. Performance: An actor, fitted and calibrated to the motion capture device with sensors and markers, performs the motions. The motion, such as body joint and facial configurations, is recorded.
3. Data Analysis: The captured data is optimized in post-production.

Motion capture has been widely used to track human and animal movements for computer character animation [4, 20, 21] because subtle details of a live performance can be applied to animated characters. As the power of computers has increased and graphics techniques have become more advanced, motion capture systems have become more affordable.

Enhancing the realism of 3D graphics has developed quickly in the last few years. The idea of creating new images from original photographs has ignited the imaginations of computer graphics researchers. It has driven them to develop new video-based techniques and tools that have been used for capturing and analyzing human movement [2, 24, 26]. The objective of the video-based techniques is to create new video clips through analyzing and rearranging original footage. These techniques are simple and practical methods that can generate new image sequences from a given video without a huge amount of labor and equipment. Because of this efficiency, it is expected that video-based techniques will facilitate simple analysis and synthesis of human and animal motion. These new technique are believed to embody efficient algorithms and tools that have many advantages over traditional motion studies such as Muybridge and Marey's photographic images.

# CHAPTER II

# RELATED WORK

In recent years, video-based techniques have been developed to generate new videos from sample video footage using appropriate combinations such as image warping and blending [8]. These simple techniques became increasingly popular because they showed the potential of animation using natural motion and photorealistic appearance as a basis.

## II.1. Video Speech Animation

To generate facial animation, a lip-stitching technique has been developed. This technique eliminates lip deformations between stitched videos to create realistic speech animation. Video Rewrite is one approach to image-based generation of character animation [2]. This facial animation system driven by audio input introduced automatic labeling of phonemes in an audio sequence. The system tracks points on the speaker's mouth from the original video footage, then reorders mouth images to go along with the phoneme sequence of the new audio track. The mouth images are then stitched into the background footage using a morphing technique. Video Rewrite demonstrates the quality of natural facial animation and the possibility of movie dubbing to sync an actor's lip motions to a soundtrack. This technique can be applied to dubbing movies, teleconferencing and special effects.

Later, Ezzat et al. presented a mouth trajectory synthesis module [11]. It is capable of synthesizing the human subject's mouth not recorded in the original video. The synthesized utterance is composited onto a background sequence including natural head and eye movement. The final output footage resembles the recorded speech of the subject. This video-realistic speech animation technique has enabled a high level of video realism: correct motions, dynamics, and coarticulation effects.

## II.2. Video Motion Capture

Video Motion Capture is a new motion tracking technique for full body movements [3, 4, 21]. It does not require markers, body costume, or other devices attached to the subject. Only the video of the person in motion is needed. Bregler et al. introduced the use of a new mathematical technique, the product of exponential maps and twist motions, and their integration into differential motion estimation. This enables a differential method based on solving linear systems that seek to recover the kinematic degrees-of-freedom (DOF) in noise and complex self-occluded configurations. The most important contribution is a multi-level, highly accurate, probabilistic architecture of human body movements. The system tracks the multiple degrees of freedom of articulated human body configuration from frame-to-frame in complex video sequences.

One interesting experiment resulted from the historic chronophotography by Muybridge. This technique was applied to recover and animate photographic images of Muybridge's motion studies from the last century. The system tracked the left and the right side of the person from multi-view sources of his photographs. They modeled the hip, shoulder, knee, elbow joints, and the neck point by DOFs and animated the tracking results successfully.

## II.3. Video Texture

Another significant technique is Video Textures, which creates a new video loop by synthesizing and rearranging the order of the original video frames [24]. It provides a continuous stream of video frames by minimizing visual artifacts so that the new video could be infinitely looped. The set of transitions in the new video can jump forwards or backwards without visual discontinuity, so the final loop is continuous and infinite. In Figure 13, the red arcs of a candle flame show the transition from one frame in the original clip to another.

The original video sequence is input into the system to analyze similarities and the system finds transition points where the video can be looped back onto itself. During synthesis, those transitions are synthesized to find a set of transitions that can create a single video loop of a given length. Finally, the final output sequence is rendered to generate a video loop, which can be played indefinitely.

Schödl et al. also presented several extensions of video textures to show diverse approaches. For example, sound samples were associated with video texture frames and re-rendered. Secondly, video textures can also be combined with image-based rendering such as view interpolation to simulate 3D motion. As a result, a 3D portrait of a smiling woman was created by mapping three videos, which were taken at different viewing angles about 20 degrees apart. Motion factorization was also one of the extensions that divided the original video into independent regions of moving parts, either manually or automatically.

In addition, mouse-controlled fish in a fish tank were created using a complex video consisting of several elements under a user or program's control (see Figure 14). The artificial fish sprites were generated by a video texture system and the tank included two sets of bubbles and swaying trees to complete the look of a fish tank. This extended technique had significant impact on many researchers and showed possibilities of building complex scenes with high accuracy.

Fig. 13. Generated video texture of candle flames. Reprinted with permission from [24]

Fig. 14. The final example of a complete fish tank. Reprinted with permission from [24]

## II.4. Video Sprite Animation

Video-based animation is a new approach in computer graphics extended to rendering and animation of realistic scenes by analyzing short video sequences. Schödl et al. focused on the video sprite, a special type of video texture [25]. They extracted characters from the green background of reordered sequences and represented them as video sprites that could be rendered at arbitrary image locations.

For manageable optimization, the technique called *beam search* was presented to control sprite motion while preserving the visual smoothness of the sequences. Schödl et al. also introduced a technique called *Q-learning* to control the sprite motion interactively for computer games. In Figure 15, original animal footage was extracted from the background using background subtraction and chroma keying, then it was corrected for size, varying with depth and perspective distortion. The fast rejection technique for computing video sprite transitions was used and a cost function was defined to generate animation. Finally, the video sprites were composited onto the desired background. Examples of a hamster and a fly in Figure 16 show the video sprite pipeline and more practically generated animation.

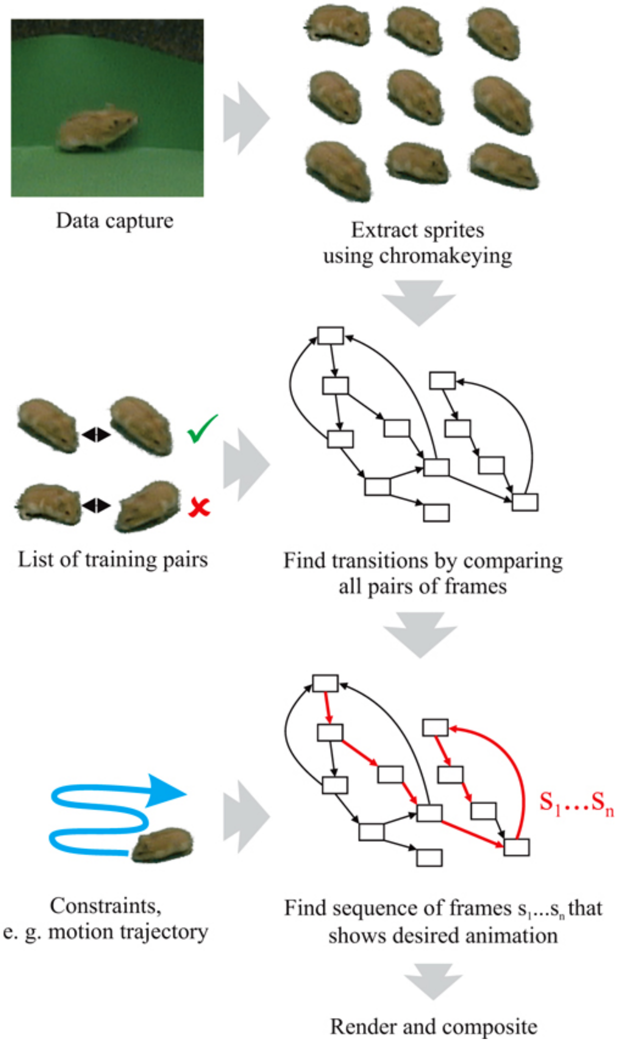


Fig. 15. Overview of video sprite animation pipeline. Reprinted with permission from [25]

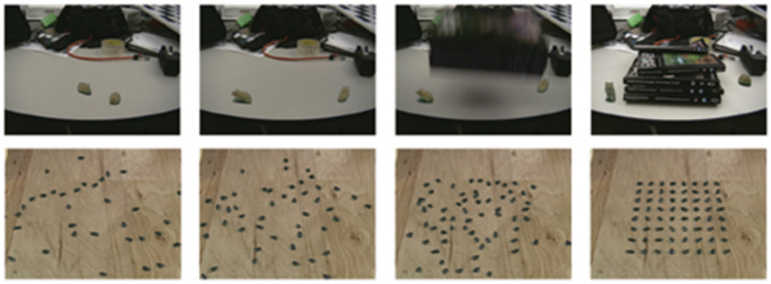


Fig. 16. The generated hamster and fly animations. Reprinted with permission from [26]

## II.5. Cartoon Capture and Retargeting

Bregler et al. recently presented cartoon capture used to track character motions from 2D classic animation in order to retarget them onto other media such as 3D models, 2D drawings, and photographs [5]. They described the challenges of two different types of capture: cartoon contours and unlabeled video. The cartoon capture method using a combination of affine transformation and key-shape interpolation showed the possibility of generating complex motion.

The idea of cartoon capture is compelling and intriguing that it captures 2D classical animations and retargets them to cartoon characters. Since motion capture using earlier computer graphics technology was limited to creating a realistic motion style, cartoon capture is very exaggerated and stylized. It allows users to study the cartoon style of characters from Disney or Warner Brother’s famous 2D animations.

# CHAPTER III

# METHODOLOGY

The main concept in Video Looping is to identify all the frames that are visually similar to each other in a given original video sequence. Using these frames, it is possible to create smooth transitions without causing significant discontinuity. These output images can be organized to create a cyclic loop.

It is easy to find identical frames in a video sequence in a simple case, such as a candle flame. Since the flame does not move from one place to another and only the shape of the flame is changed, color values of the input frames can be simply compared to find similarities. On the other hand, human or animal motions are not straightforward in terms of distinguishing color values between input frames. Computing color values of video images is difficult because subjects do not stay in the same position- they move constantly. In order to simplify this comparison, original video footage was shot in front of the bluescreen background with a stationary camera. For the purpose of extracting a moving character from the background, a bluescreening technique was required. The following is an overview of the Video Looping system in Figure 17.

## III.1. Shooting Video

The videos were primarily shot in front of the bluescreen at a studio to separate the moving character from the background. A stationary camera was locked in one position and placed perpendicular with the moving character in order not to show shadows on the floor. Additionally, the character must always be in motion to create a natural cyclic loop.

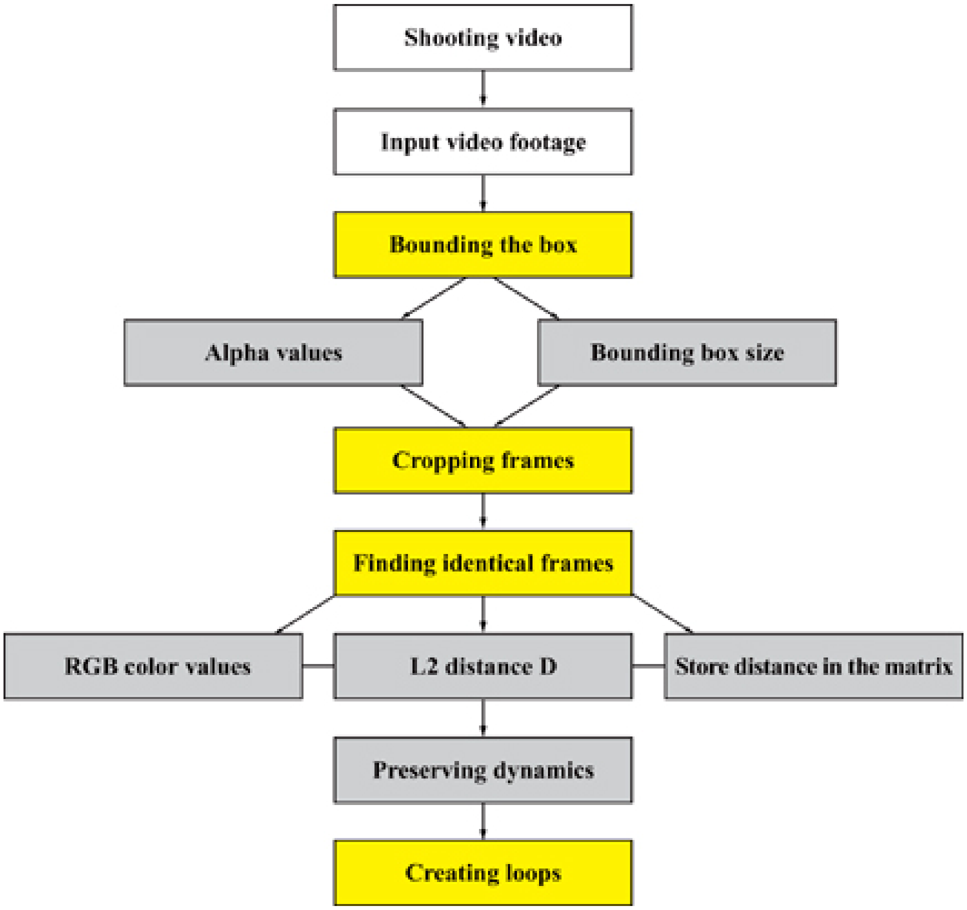


Fig. 17. System overview diagram.

## III.2. Bounding the Box

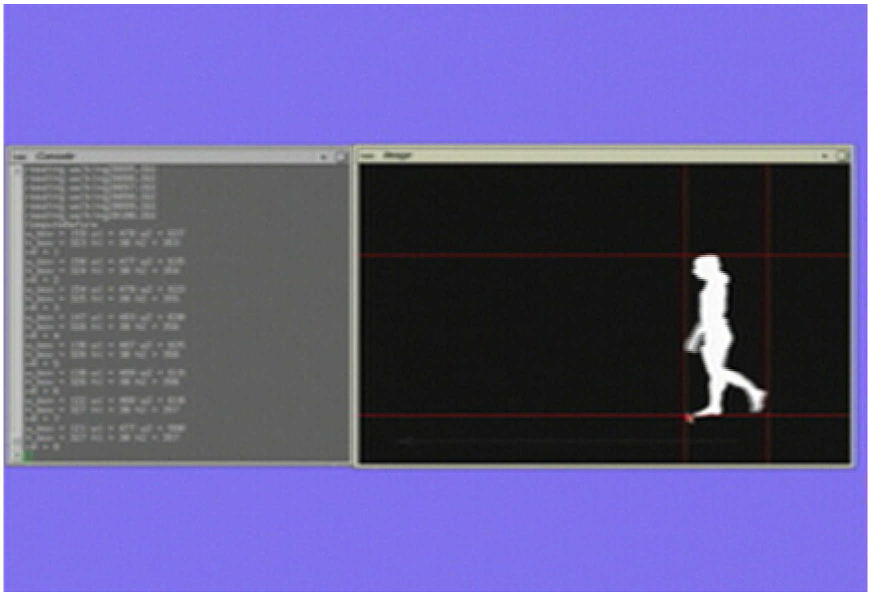
The bounding box is the smallest moving rectangle that includes the moving character. This rectangle guarantees that the character always stays inside of the box. Technically, it represents the translation of a moving subject in each frame. It tracks the position of the character and compares translation values and size of the box in each frame.

To determine the smallest rectangle that contains the character in motion, alpha values were computed using a bluescreen technique. As shown in Figure 18, the maximum value of the bounding box was chosen as the output image size and it became a guide for cropping all output images.

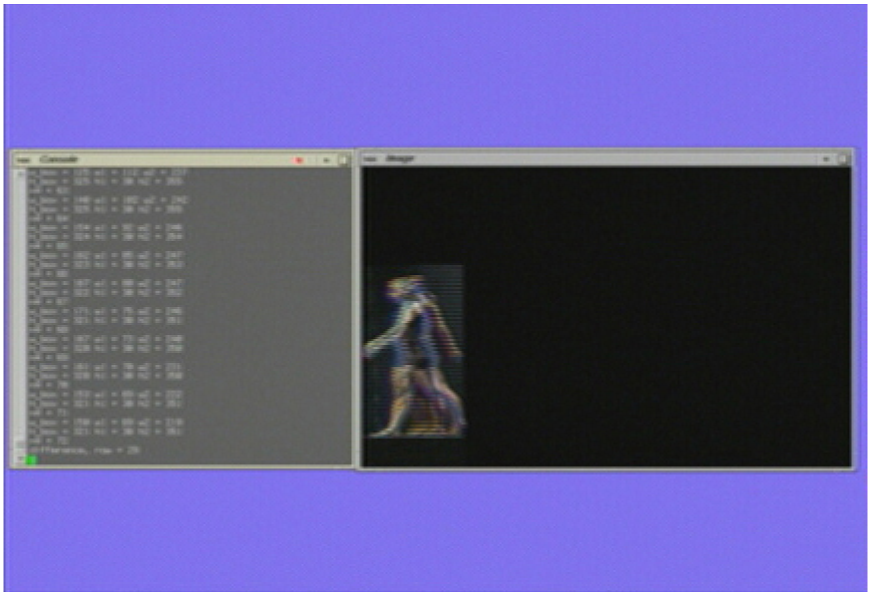
## III.3. Cropping Frames

As shown in Figure 22, based on the bounding box, the resulting images were first cropped and placed then in the same position. Because the character's positions were different in each frame, they needed to be cropped in order to create a natural cyclic motion. The final loop looked like a natural cycle through placement of output images in the same position. An additional margin was added to the maximum size of the box for the output images. Then the final output images were translated for looping (see Figure 19).

Cropping provides a set of images in which the feet and the vertical center of the character's bounding box are exactly aligned. After this alignment, one can observe the character walking or running in the equivalent position.



(a)



(b)

Fig. 18. Compute alpha and RGB values using Video Looping system. (a) Computing alpha values and the bounding box size of a character. (b) Calculating color values.

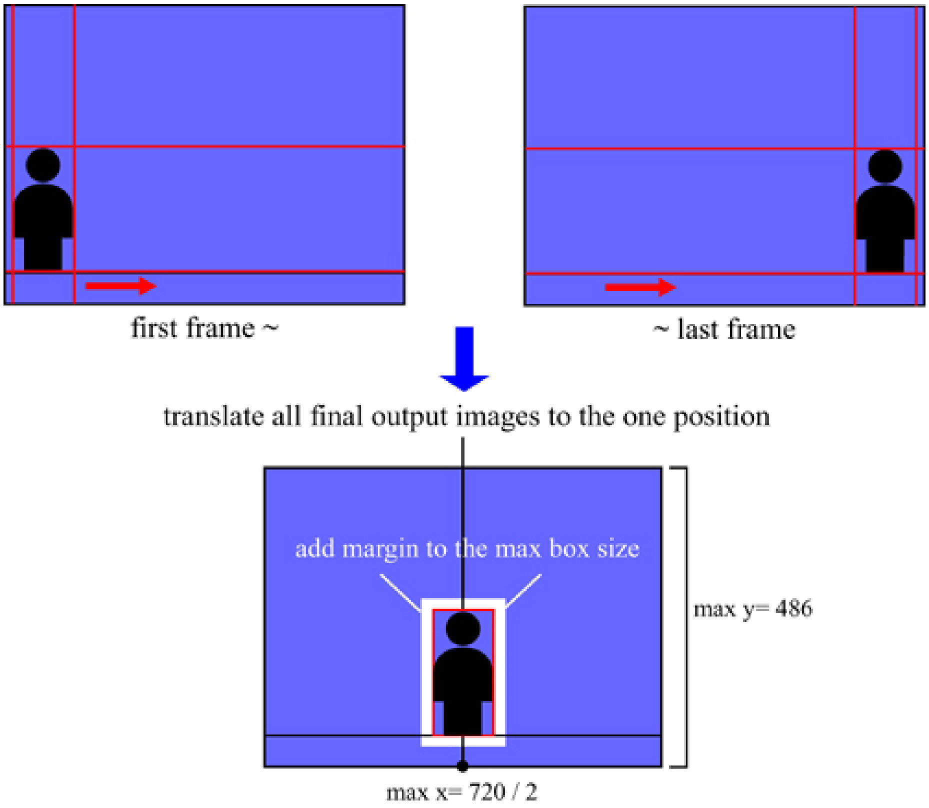


Fig. 19. Cropping final output images based on the bounding box.

## III.4. Finding Identical Frames

In this step, input video was analyzed to find the best transition points for looping without visual discontinuity. Figure 20 shows how such a loop was created with a smooth transition from frame $i$ to $j$. In a transition from sample i to sample $j$, the successor of sample $i$ by sample $j$ can be substituted the predecessor of sample $j$ by sample $i$. It means that frame $i$ should be identical to frame $j-1$ and frame $j$ should be identical to frame $j+1$. Otherwise there could be noticeable artifacts. If these conditions were satisfied, the resulting cyclic loop can repeat indefinitely.

The system computed the sum of the difference of RGB color values at every pixel of the two different frames. It is for the purpose of finding the similarity between two images, which is represented by distance function ($D_{ij}$). Similar to the analysis step of video texture, comparison was done using $L_2$ distance [24]. Once all the frame-to-frame distances were computed, the system stored them in the matrix, $D_{ij}$. Transitions from frame $i$ to $j$ can be performed when $D_{i+1,j}$ is small. The system stored the probability of these transitions in a probability matrix and normalized each transition [24]. The smallest values of distance function corresponded to the most identical frames in the sequence. These frames were good candidates for the smoothest transitions.

Although frame $i$ to $j-1$ may create a good transition (see Figure 20), these two frames may sometimes represent opposite motions. In the example of a pendulum sequence shown in Figure 21, the pendulum is swinging from left to right in the top row. However, the center frame matches both frames in the bottom row. While the $i$ sequence shows the same left to right motion like the $j_1$ sequence, the $j_2$ sequence shows a pendulum swing from right to left, so the transition suddenly reverses the motion. In order to avoid such abrupt changes in motion, the dynamics of motion needed to be preserved by comparing not only all frames themselves, but also small subsequences of adjacent frames [24]. In other words, the frames $i + m$ should be similar to the frames $j + m$ for smooth transition from frame $i$ to $j$.

### III.5. Creating Loop

Once similarities were identified, the system chose the minimum number of frames, which can make the best transitions to create a final loop. It displays the number of frames that are chosen and the total frame's range for a loop. As shown in Figure 22, the final loop was continuously played using OpenGL.

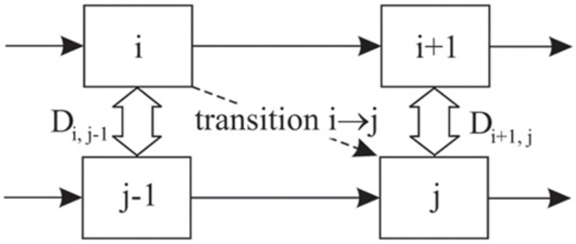


Fig. 20. Image similarities and transitions. $D_{ij}$ is the distance function. Reprinted with permission from [25]

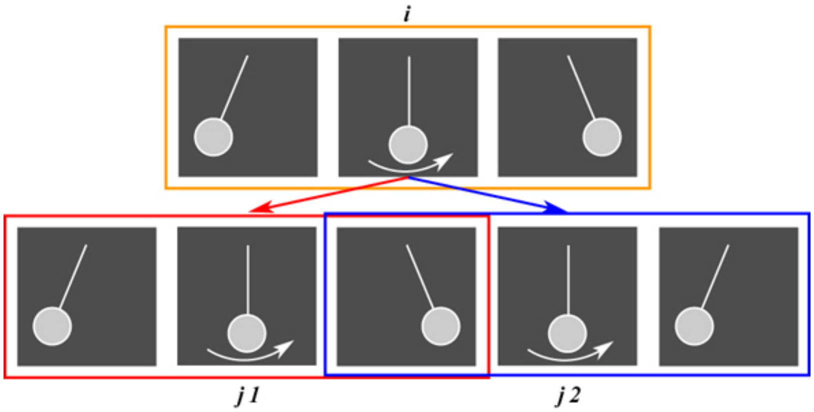
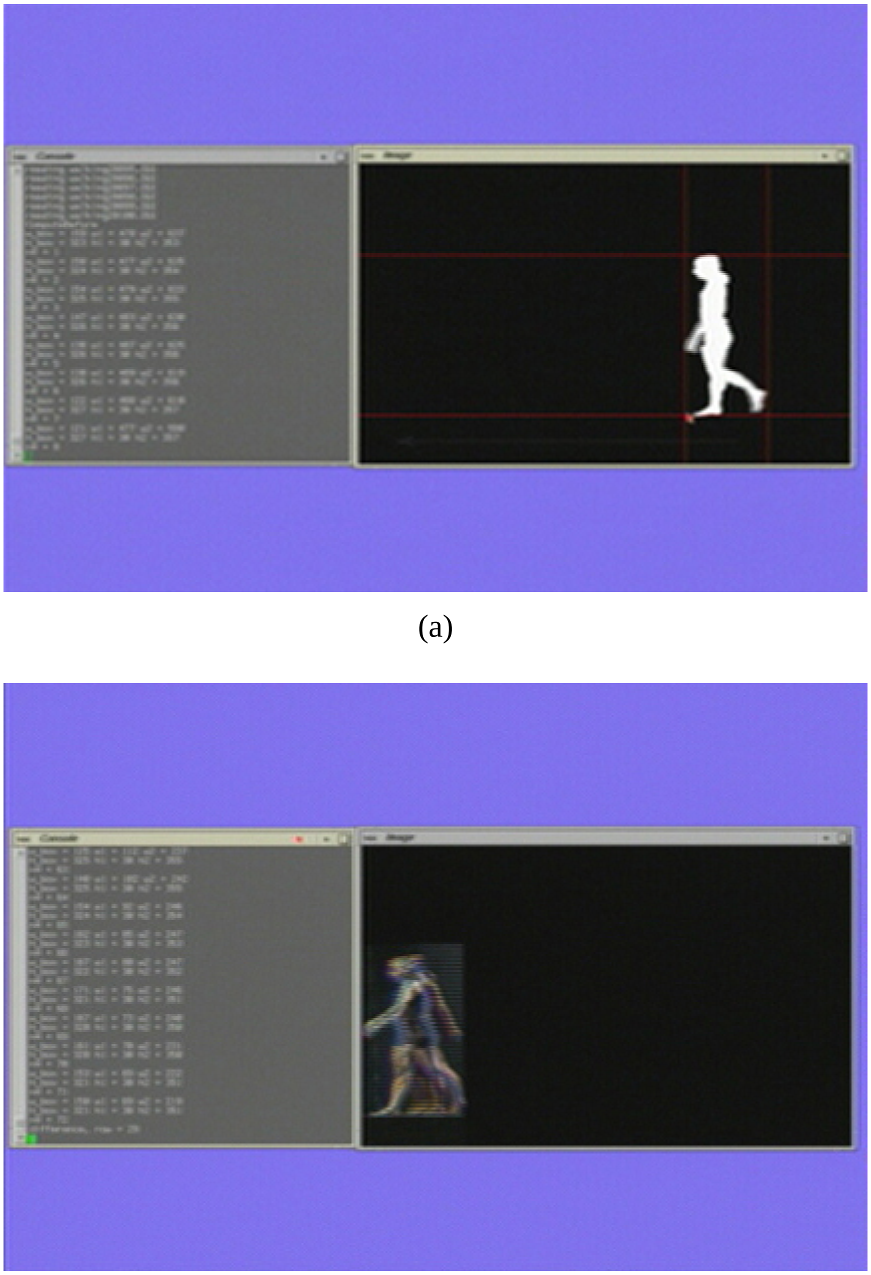


Fig. 21. Finding good transitions in a pendulum sequence.



(a)

(b)

Fig. 22. Captured screen shots of Video Looping program in Irix. (a) Probability matrix. (b) Play of a final cyclic loop using OpenGL.

# CHAPTER IV

# IMPLEMENTATION AND RESULTS

## IV.1. Video Looping Tool

To show the effectiveness of Video Looping, a stand-alone system was implemented using C++, together with OpenGL [22] and the Image Format Library$^{TM}$ (IFL). It currently runs on both SGI and Windows platforms.

Inspired by Muybridge's motion studies, three major classes - male, female and child motion - were chosen. For each class, walk, run, hop, jump cycles, and carrying a heavy object were created as a proof-of-concept. An animal case was experimentally added to demonstrate the flexibility of the system.

Each character's consecutive movements were recorded. The characters performed motions on a flat ground plane. Recording the subject in perspective view created one more step: correction for size, varying with depth, and perspective distortion. Shooting the video laterally showed leg and arm motion precisely, and the recorded images did not need to be re-projected onto a virtual plane. The recorded footage was digitized and file sequences were inputted to the Video Looping system. The IFL by Silicon Graphics Computer Systems was used to read and write image files in a format-independent fashion.

As the fixed camera displayed limited views, and each person's gait, stride and velocity was different, so the entire sequence of recording video included three to seven cycles. In general, a few frames at the beginning and the end of input sequences were not good candidates to generate a loop since part of character's motion was not entirely recorded. Several frames at the beginning and end the sequence can be regarded as garbage frames. They were excluded to save time when the program ran.

The program accurately distinguishes all moving subjects in the scene; even subtle movements, like leaves blown by wind, can become part of the bounding box. Therefore, the shadows of a character on the floor or flickering shadows on the edge of the bluescreen by studio lights were included in the bounding box. In order to minimize shadow problems, the camera was set up perpendicular to the moving character. In the program, initial margins to exclude shadows on the fringe of the bluescreen were added, and the program then checked moving subjects inside of these margins.

After the analysis of similarities, the final output images that would make a good transition were selected. The program computed minimum frames for a cyclic loop from hundreds of input frames, and approximately 10 to 40 output images were created. The system can change the minimum range of a cycle, increase it to get more than one cycle, and create two looping cycles without visual artifacts. Additionally, because a character was in motion, the background of each output frame was changed, making it difficult to judge whether the transitions in a final loop were smooth. But the program can easily delete the background using the bluescreening technique.

## IV.2. Digital Cyclic Motion Library

The purpose of a digital cyclic motion library is to help users create elaborate animations. It is also an on-line learning gallery [6], so that users can study details of human cyclic motion and download sequences and movie files (see Figure 23).

The sequences in this library include representatives of human cyclic motions which were analyzed by the Video Looping system. Final output images were arranged in order to demonstrate progressive motion and QuickTime movie files were made to loop them continuously.

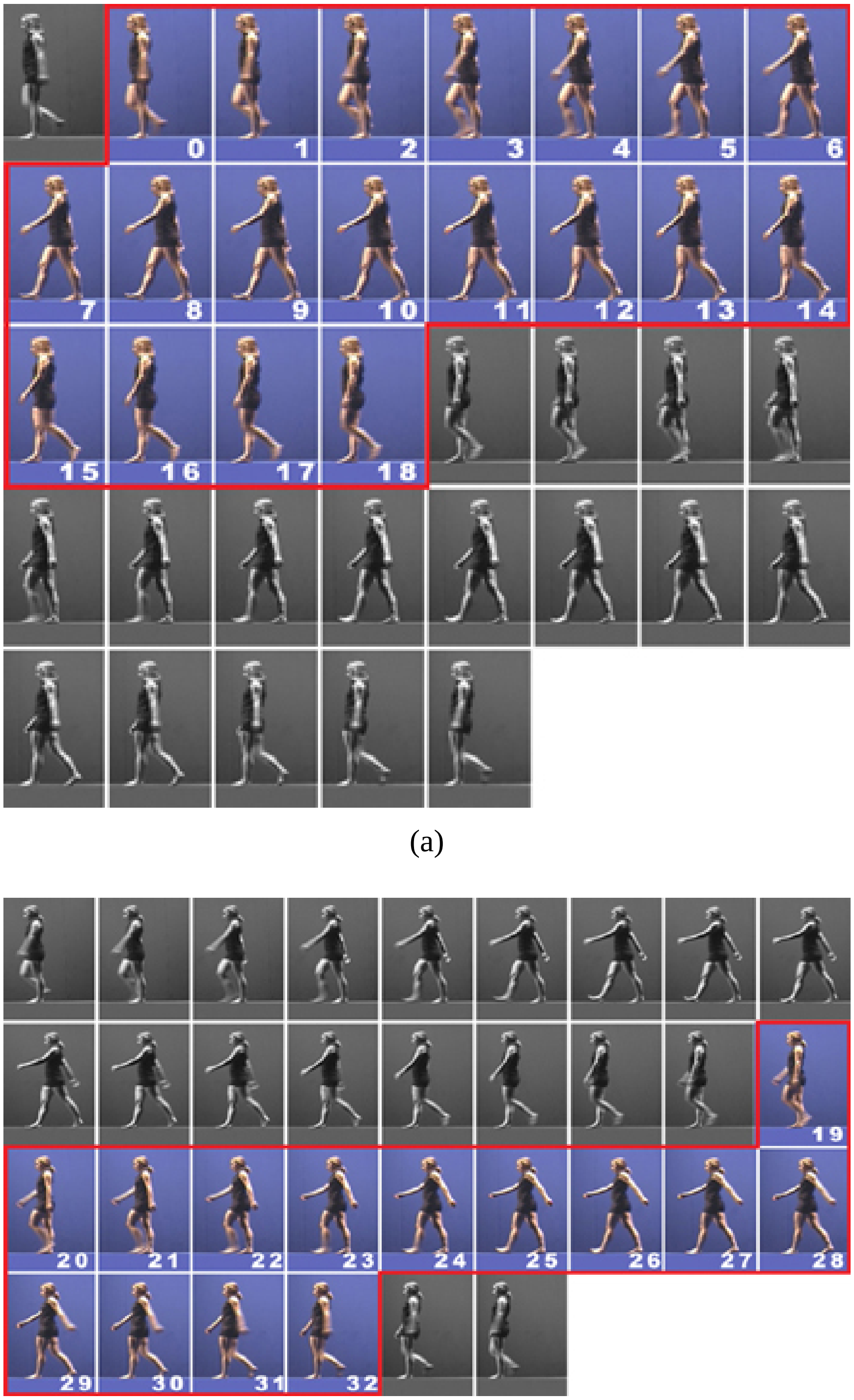
Each figure clearly demonstrates consecutive movement of a male, female and a child that resembles Muybridge's photographic analysis. A few sequences were turned into GIF animations of small size. This can improve web surfing for online users by providing faster access to web pages. Due to this considerable reduction of file size, the Vide Looping system can be effectively used to create web animation clips and present ideas in ways that were not previously possible such as making small clips from larger video files. Final GIF animations and image sequences can be downloaded on the visualization laboratory web site (http://www-viz.tamu.edu/faculty/ergun/research/motioncapture/library/).

In addition, two different loops of the same cyclic motion were combined together as shown in Figure 24. All frames of the final loop of Walk 1 and Walk 2 were compared and frames from the first loop were randomly chosen. In the final sequence of Walk 1, frame 1 to 19 were renumbered 0 to 18 and frame 18 to 31 from Walk 2 were chosen and renumbered 19 to 32 by the same manner. The renumbered frames from Walk 2 were placed after the last frame of Walk 1. The final cyclic loop was a good result without visual artifacts.

The following classes show a noteworthy confirmation of the system in each case. In each figure, it reveals the details of cyclic motions.

Input: 90, Best frames: 59-21, Range: 38

Fig. 23. Captured screen images of Digital Cyclic Motion Library.



(a)

(b)

Fig. 24. Combining two different final loops by skipping some frames. (a) Select frame 1-20 from the final loop of Walk 1. (b) Select frame 18-31 from the final loop of Walk 2.

### *IV.2.1. Walk*

As the walk cycle was Muybridge’s first consideration for animal motion, it was also the main focus of this thesis. Most of the cycles of male, female, child, and dog (see Figure 25 and Figure 26) were good natural motion without visual artifacts. In normal walking, the arms always oppose the legs to give balance. When the character’s leg is straight up, the body, head and pelvis are slightly higher than the leg as shown in Figure 27 and Figure 28. The arm swing is widest when the character’s leg is bent and it takes the weight. Although each cycle of a healthy child was a different stride and an irregular velocity, the final loops show satisfactory results.

While animals such as a cat and dog were enjoying more freedom of movement, they were more difficult to work with. Unexpected motions, like leaping or jumping, were accidental interruptions in regular cyclic motions. It is desirable to shoot the video in a large studio in front of a blue or green screen.

Fig. 25. Dog walk. Total frames: 36, best frames: 14-6, range: 8.

Fig. 26. Child walk. Total frames: 206, best frames: 139-119, range: 20.

Fig. 27. Female walk. Total frames: 131, best frames: 88-55, range: 33.

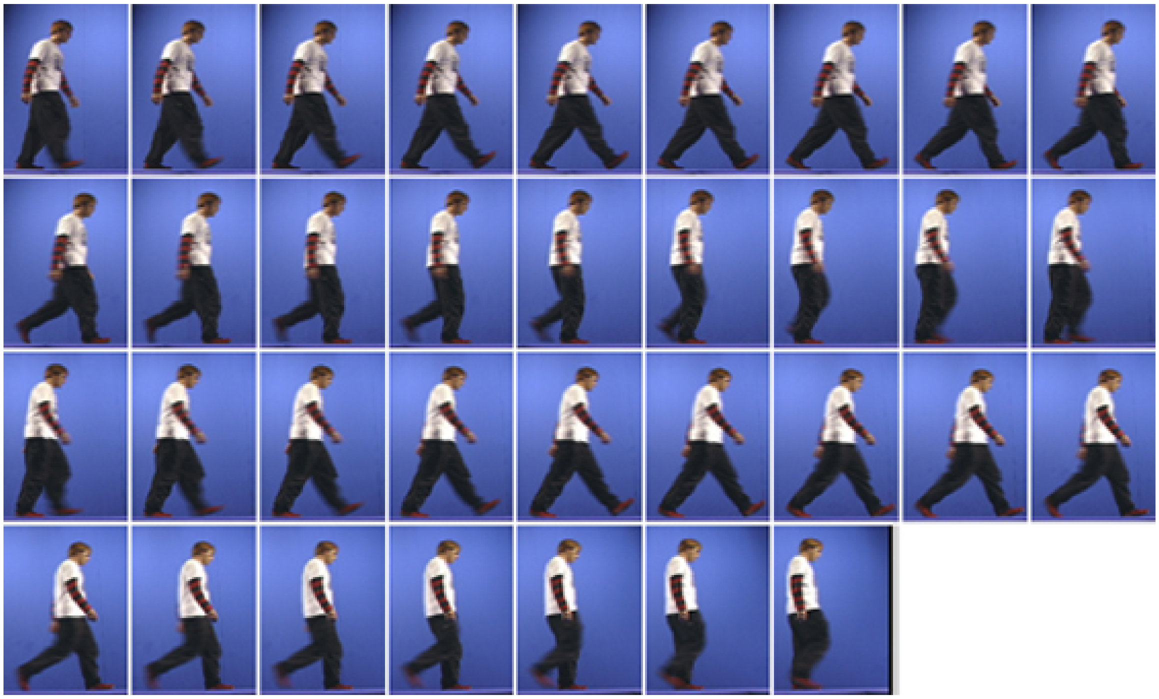


Fig. 28. Male walk. Total frames: 70, best frames: 43-9, range: 34.

### *IV.2.2. Run*

Running can be regarded as fast walking because a subject performs almost the same motion in approximately half the time. The difference between a walk and a run is in the position of the foot. In a walk cycle, one foot of a character is always on the ground and the other is off the ground. As shown in Figure 29, Figure 30, and Figure 31, both feet of a character are off the ground in some frames. The arm motion is more restricted than a walk and it does not swing as much.

To record a run cycle, the camera needed to be set up further from the character, but there was a distance limit in the studio. Therefore, the input video was not a run at full speed but was instead a jog at regular speed. A jog is slightly slower than a standard run because the length of the strides is much shorter. The head does not pump up and down as in a fast run. In a fast run, the head and body usually are more forward and the arm movement is faster. The following images show the output loops of the running motion.

Fig. 29. Child run. Total frames: 77, best frames: 24-9, range: 15.

Fig. 30. Female run. Total frames: 161, best frames: 109-80, range: 29.

Fig. 31. Male run. Total frames: 131, best frames: 90-67, range: 23.

### *IV.2.3. Hop and Jump*

Other types of interesting cyclic motions are hopping and jumping. There can be a large variety in these motions such as double bouncing and skipping. In the case of a hop with two legs (see Figure 32 and Figure 33), the motion is similar to a jump. In Figure 34, a hop cycle with one leg, a character goes down on one foot and hops on the same foot. The output images of both jump and hop show where the ups and the downs are on the different parts of the body. Timing is different in each output loop when a character was in the air.

If a character jumps in the air, the arms, feet, and body of a character is extremely stretched for few frames. The whole body is then squashed when the character lands on the ground. The spine is bent, and reverses on landing. There is a similarity to the run in Figure 35 and Figure 36. Between each jump cycle, there is a short skip and an anticipation of next jump. Within these frames, the arms are swinging wide to prepare for the motion. The jump cycle is not fully athletic as with jumping a hurdle.

Fig. 32. Male hop. Total frames: 145, best frames: 79-61, range: 18.

Fig. 33. Female hop. Total frames: 79, best frames: 34-18, range: 16.

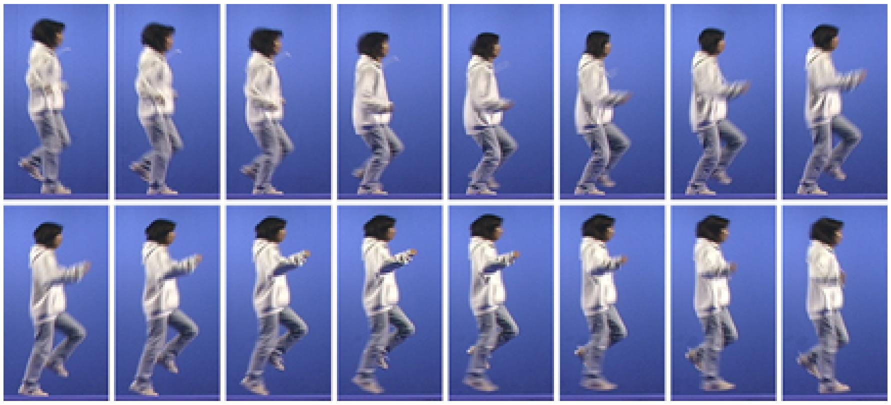


Fig. 34. Female hop with one leg. Total frames: 113, best frames: 58-42, range: 16.

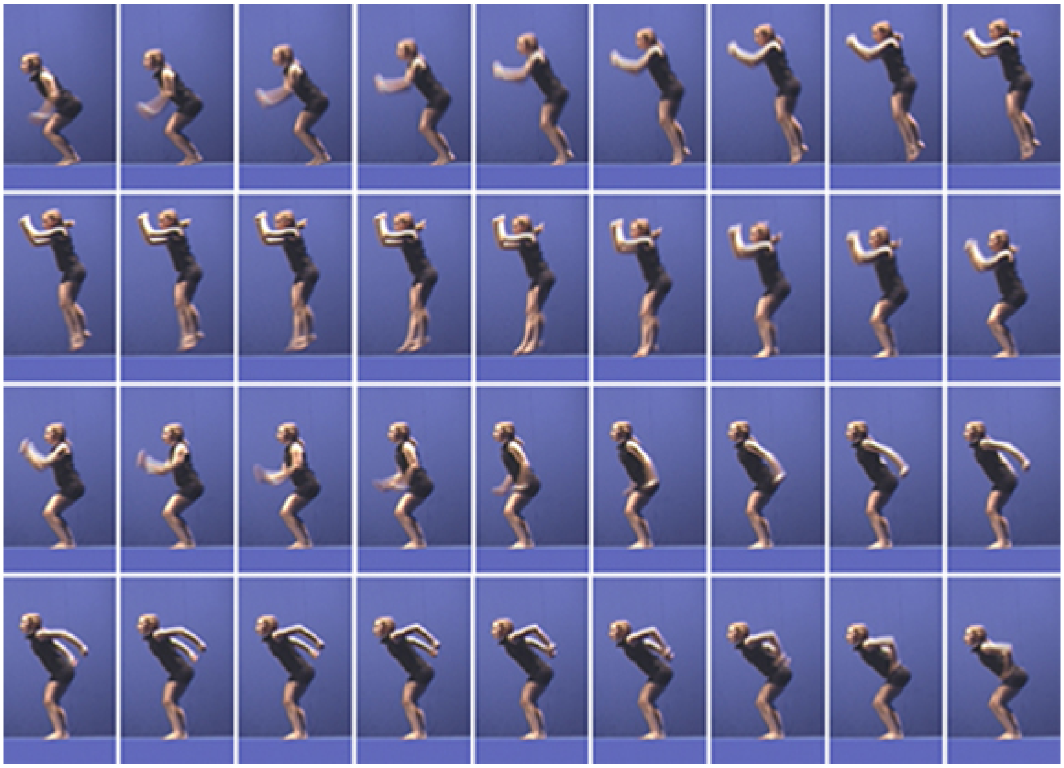


Fig. 35. Female jump. Total frames: 296, best frames: 102-66, range: 36.

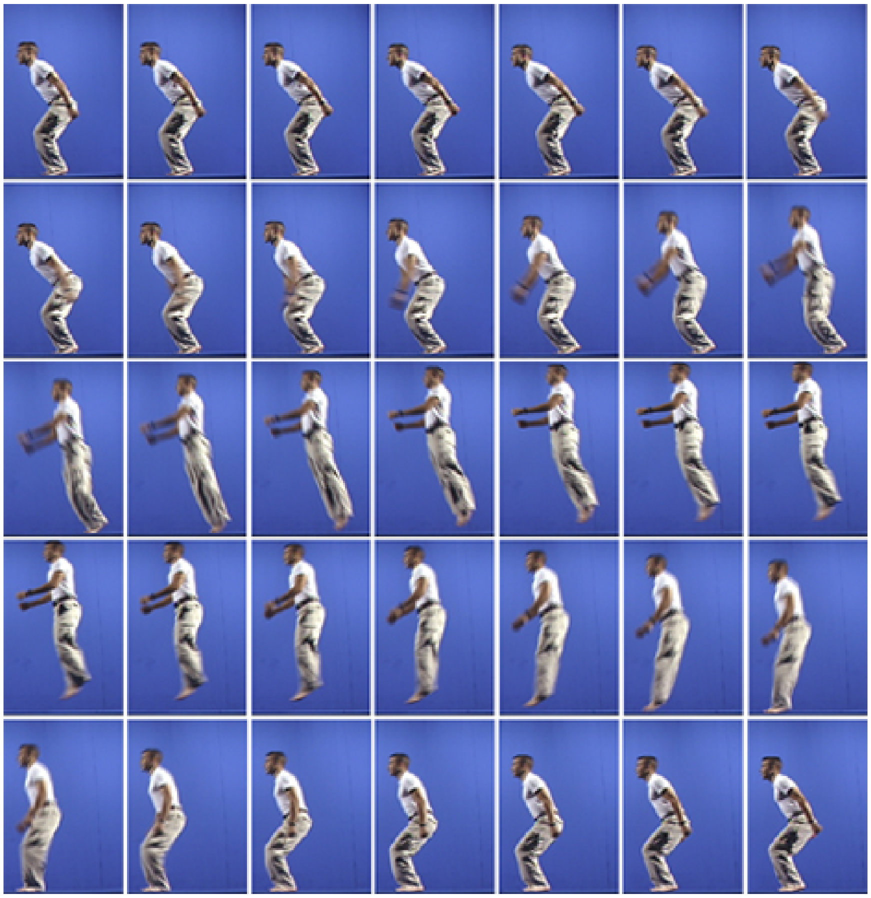


Fig. 36. Male jump. Total frames: 206, best frames: 119-74, range: 35.

### *IV.2.4. Carrying a Heavy Object*

Carrying a heavy object can precisely show where the weight on a character is at each frame. Spreading feet, bending knees and spine, back arch, and changes in motion reveal the body's balance (see Figure 37 and Figure 38). In input video, there was no anticipation or adjustment before a character lifted the object, but Figure 37 shows well how the character controlled her weight to carry it. The weight sinks with the shoulders, and arms are continuously stretched. The knees are bent all the time and the feet do not come off the ground as much as in a walk. The object is not raised although the upper body is moved slightly up and down. Sometimes, the timing of the feet could be irregular because the character steps and pauses to balance the weight of the object.

Fig. 37. Female carrying an object. Total frames: 143, best frames: 59-32, range: 27.

Fig. 38. Male carrying an object. Total frames: 112, best frames: 50-14, range: 36.

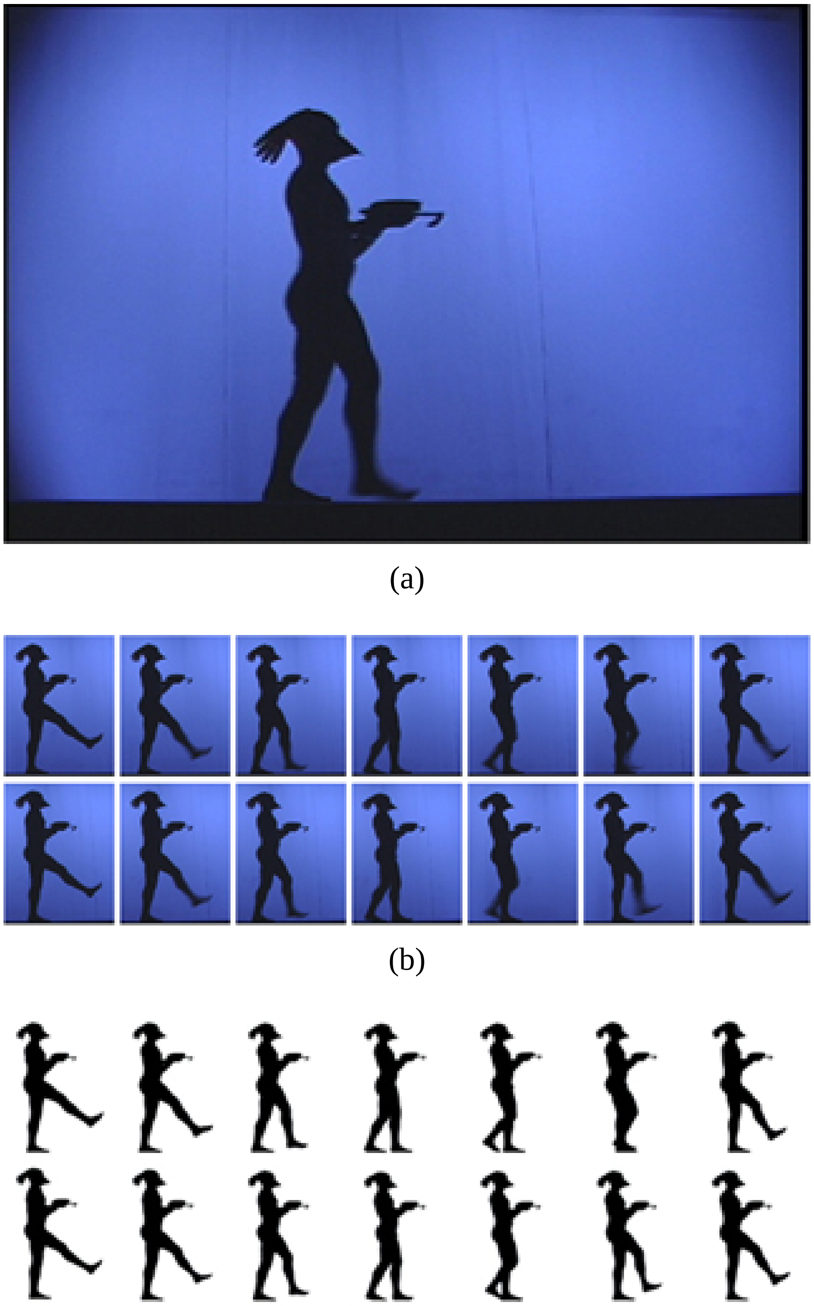
### IV.3. Compositing Multiple Video Loops

Video sprites, in which video textures are rearranged and re-rendered after background subtraction [24, 25, 26], have frequently been used to create animation. Although a video sprite's value is judged by the ease with which a complicated scene can be rendered, duplicating video loops can also become an interesting solution. Since the output images can be easily separated from the bluescreen background (see Figure 39), a video loop can be duplicated and composited onto different backgrounds to create a complex scene.

A 2D animation, which was composed of several video loops, was created using a Video Looping system. In this animation, a cyclic loop of a man walking was generated. Originally, the output frames were 40 frames (frame 126-86), but half were skipped to make a stylized motion. Only fourteen frames were used to create a cyclic loop and the character was extracted from the bluescreen background as shown in Figure 40.

Various background illustrations and other visual elements were drawn using commercial graphics software such as Painter® and Adobe Photoshop®. All visual elements were then imported into motion graphics software Adobe After Effects® and were set up in different layers for animation. The character's cyclic loops were composited onto the background illustration (see Figure 40). Each layer for the background and other additional visual elements was key-framed in After Effects®. Figure 39 and Figure 40 shows the process of creating an experimental animation with composited multiple cyclic loops.

The final animation can be downloaded and played on the visualization laboratory web site (http://www-viz.tamu.edu/faculty/ergun/research/motioncapture/library/project).

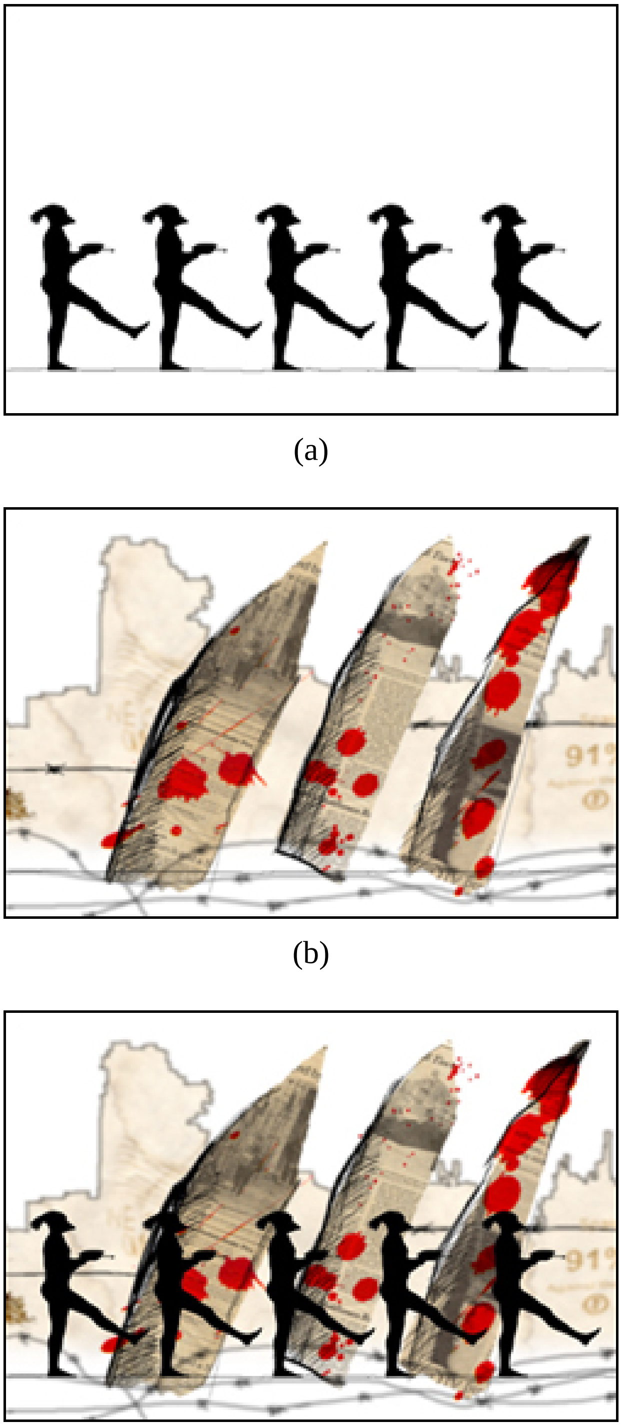


(a)

(b)

(c)

Fig. 39. Shooting and bluescreening technique. (a) Input frames with the bluescreen: 0-230. (b) Final skipped frames: range 14. (c) Extracted the character from the bluescreen.



(a)

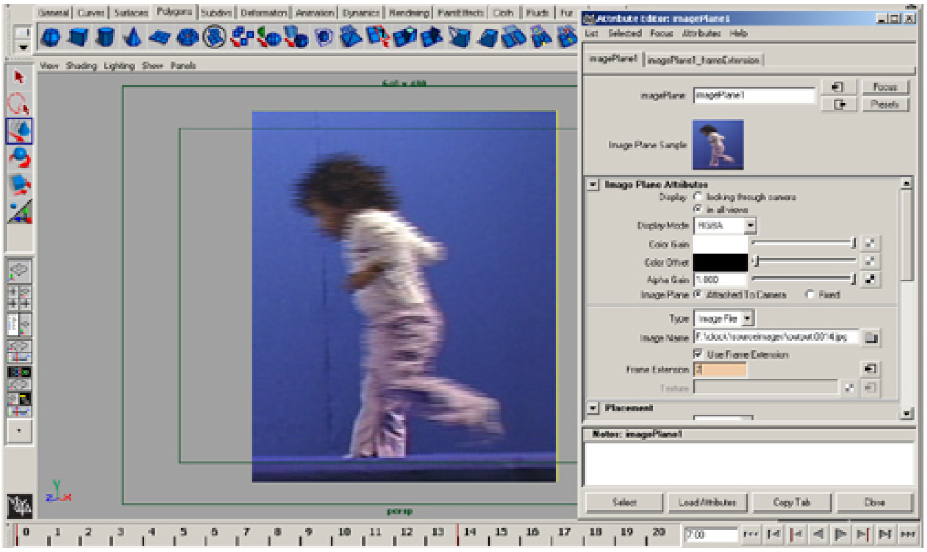
(b)

(c)

Fig. 40. The process of creating 2d animation, ’March of the Chicken-Man’. (a) Duplicating a loop of character’s walk cycle. (b) Digital illustration for background. (c) Final composited new shot.

## IV.4. Rotoscoping in 3D Animation

Traditional animators have used rotoscoping to build elaborate sequences and to establish a basis of movement and rhythm. Rotoscoping is the tracing of a photographic image, projected onto animation paper or an animation cel, to create a series of animated frames. The final video loops, which are generated from the Video Looping system, are good references to animate a character; therefore they can be used as a template for 3D character animation. As shown in Figure 41, the image sequence of a cyclic loop was imported to the 3D software Maya® and users can then create keyframe animation using the rotoscoping technique. As the final output images were from a cyclic loop that did not have discontinuity, it saved time to tweak the animation curves. This technique simplifies keys to achieve smooth animation and goes further to exaggerate motion for a cartoon character style.



imagePlane: imagePlane1

Display Mode: RGBA

Alpha Gain: 1.000

Type: Image File

Fig. 41. Importing the final images to 3D software, Maya®.

# CHAPTER V

# CONCLUSION AND FUTURE WORK

Analysis of human motion is still difficult and challenging because simple cyclic motion of humans contains massive amounts of information that convey a character's personality.

Video Looping is a system to generate a cyclic loop, which does not require an intricate process and special skills. It is a practical tool that compares similarities to find the point of good transition from a simple video and make a natural motion of a cyclic loop without visual discontinuities. It is also a learning tool to study the subtleties of the cyclic motion of different characters. Moreover, it is a flexible tool that can be applied to various experimental and artistic multimedia projects.

The digital cyclic motion library contains representative figures as well as Muybridge's and Marey's photographic analyses. It is an authentic and clear resource for tracking the detail of human cyclic motions and creating keyframe animations. Furthermore, cyclic motion can be used to compare differences in each character's motion, emotion, and personality. In the example of a fat man's angry walking, the weight shift from one foot to another is more pronounced, the arm movement is wilder, and the foot action faster and wider than normal. Footage can be prepared for different characteristics such as happy or sad, young or old, and fat or skinny.

Future work could be directed towards creating video-based animations using video loops. As sprites have been the most popular graphic sources for creating computer games, they can be combined with output loops to generate a more complex motion for computer games. Once a character's cyclic motions are extracted from the background, one cyclic motion can be blended into another, such as a walk to a run or a run to a jump. Those loops can be under interactive control in such a way that users are able to specify segments within a video source and speed it up or slow it down. In fact, similar enhancements are currently being investigated by Schödl *et. al* for the development of video-based animations that give controllable functionalities for users in real time [26].

Combining two different cyclic motions is relatively difficult since finding smooth transitions have been a problem. A more precise method should be developed to interpolate two motions without visual discontinuity at the point of instantaneous transition from one to another. Warping or blending techniques can disguise visual artifacts, though it is a time-consuming process. Consequently, more accurate analysis of complex motions must be accomplished together with warping and blending.

# REFERENCES

[1] M. Braun, *Picturing Time: The Work of Etienne-Jules Marey (1830-1904)*. London: The University of Chicago Press, 1992.

[2] C. Bregler, M. Covell and M. Slaney, “Video Rewrite: Driving Visual Speech with Audio”, *Proc. SIGGRAPH 97*, pp. 361-368, August 1997.

[3] C. Bregler and J. Malik, “Video Motion Capture”, Technical Report UCB//CSD-97-973, Dept. of Computer Science, University of California, Berkeley, 1997. http://www.cs.berkeley.edu/bregler/digmuy.html.

[4] C. Bregler and J. Malik, “Tracking People with Twists and Exponential Maps”, *Computer Vision and Pattern Recognition 1998*, pp. 8-15, June 1998.

[5] C. Bregler, L. Loeb. E. Chuang, and H. Deshpande, “Turning to the Masters: Motion Capturing Cartoons”, *Proc. SIGGRAPH 02*, pp. 8-15, July 2002.

[6] H. Choi, “Digital Cyclic Motion Library”, October 2003.
http://www-viz.tamu.edu/faculty/ergun/research/motioncapture/library/

[7] F. Dagognet, *Etienne-Jules Marey: A Passion for the Trace*. New York: Urzone,Inc., 1992.

[8] P. Debevec, “Image-Based Modeling, Rendering, and Lighting”, *Siggraph 1999, Course 39*, August 1999.

[9] Disney Enterprises, Inc., *Snow White and The Seven Dwarfs Platinum Edition*. Burbank: Buena Vista Home Entertainment, Inc., 2001.

[10] L. Eitner, *Eadweard Muybridge: The Stanford Years 1872-1882*. The Leland Stanford Junior University, Palo Alto, 1972.

[11] T. Ezzat, G. Geiger, T. Poggio, “Trainable Videorealistic Speech Animation”, *Proc. SIGGRAPH 02*, pp. 388-398, July 2002.

[12] M. Gleicher, “Motion Editing with Spacetime Constraints”, *Proc. Symposium on Interactive 3D Graphics 1997*, pp. 139-148, April 1997.

[13] J. Lasseter, “Principles of Traditional Animation Applied to 3D Computer Animation”, *Proc. SIGGRAPH 87*, pp. 35-44, July 1987.

[14] G. Lista, *Futurism*. Paris: Finest S.A. / Éditions Pierre Terrail, 2001.

[15] J. Lee, and S.Y. Shin, “A Hierarchical Approach to Interative Motion Editing for Human Like Figures”, *Proc. SIGGRAPH’ 99*, pp. 39-48, August 1999.

[16] L. Maltin, *Of Mice and Magic: A History of American Animated Cartoons*. New York: New American Library, 1980.

[17] E. Muybridge, *Human Figure in Motion*. London: Dover Publications, 1955.

[18] E. Muybridge, *Animals in Motion*. London: Dover Publications, 1957.

[19] E. Muybridge, *Muybridge’s Human and Animal Locomotion*. London: Dover Publications, 1957.

[20] K. Pullen, “Animating by Multi-level Sampling”, *IEEE Computer Animation Conference 2000*, pp. 36-42, May 2000.

[21] K. Pullen, “Motion Capture Assited Animation: Texturing and Synthesis”, Ph.D. Thesis, Dept. of Physics, Standford University, Stanford, July 2002.

[22] M. Woo, J. Neider, T. Davis, D. Shreiner, *OpenGL Programming Guide: The Official Guide to Learning OpenGL, Version 1.2*. Reading, Massachusetts: Addison-Wesley, 1999.

[23] M. Rush, *New Media in Late 20th-Century Art*. London: Thames & Hudson, 1999.

[24] A. Schödl, R. Szelinski, D.H. Salesin and I.A. Essa, “Video Textures”, *Proc. SIGGRAPH 00*, pp. 489-498, July 2000.

[25] A. Schödl and I.A. Essa, “Machine Learning for Video-Based Rendering”, *Advances in Neural Information Proceeding Systems*, Cambridge: MIT Press, T.K. Leen, T.G. Dietterich, and V. Tresp, eds, vol. 13, pp. 1002-1008.

[26] A. Schödl and I.A. Essa, “Controlled Animation of Video Sprites”, *Proc. SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 121-127, July 2002.

[27] C. Solomon, *Enchanted Drawings : The History of Animation*. New York: Alfred A. Knopf, Inc., 1989.

[28] B. Thomas, *Walt Disney: An American Original*. New York: Simon and Schuster, 1976.

[29] B. Thomas, *Disney’s Art of Animation : From Mickey Mouse to Beauty and the Beast*. New York: Hyperion Press, 1991.

[30] F. Thomas, O. Johnston. and C. Johnston, *The Illusion of Life: Disney Animation*. New York: Hyperion Press, 1995.

# APPENDIX 1

# EXTRA WORK

In the case in which the background is not a bluescreen background, the bounding box was determined based on a character's motion. Since the color values of a character and the background were similar, color values were obtained by subtracting RGB values of frame $i$ and frame $i + 1$. As shown in Figure 42, maximum and minimum values of x and y were found and they represented each boundary of the bounding box.

To find frame $j$, which was not overlapped to the bounding box of frame $i$, temporal differences of RGB values were used. Simply, correct color values can be obtained by comparison between frame and frame, which are not overlapped. In the frames, which were from the first to less than half of the input frames, the motion of a character did not overlap with the last frame, so color values could be correctly obtained. On the other hand, the remainder of the frames was compared with the first frame of input sequences. If the $L_2$ distance between frame $i$ and the first or last frame was greater than 0, the alpha value was also greater than 0. If $L_2$ distance was less than 0 or equal to 0, then the alpha value also was 0 (see Figure 43).

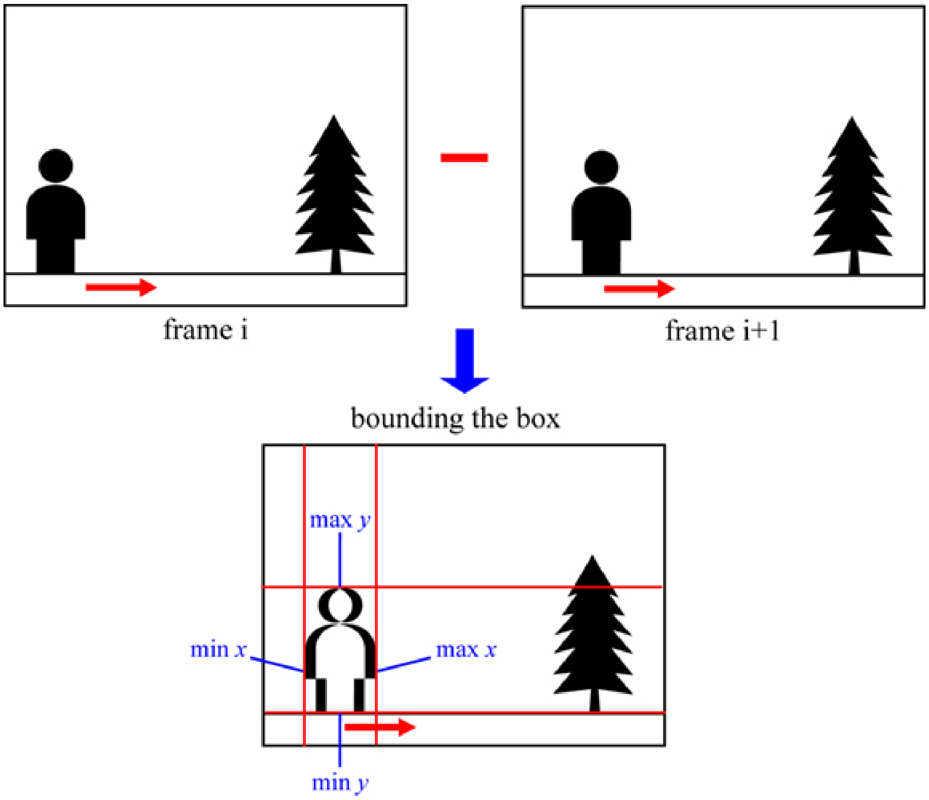
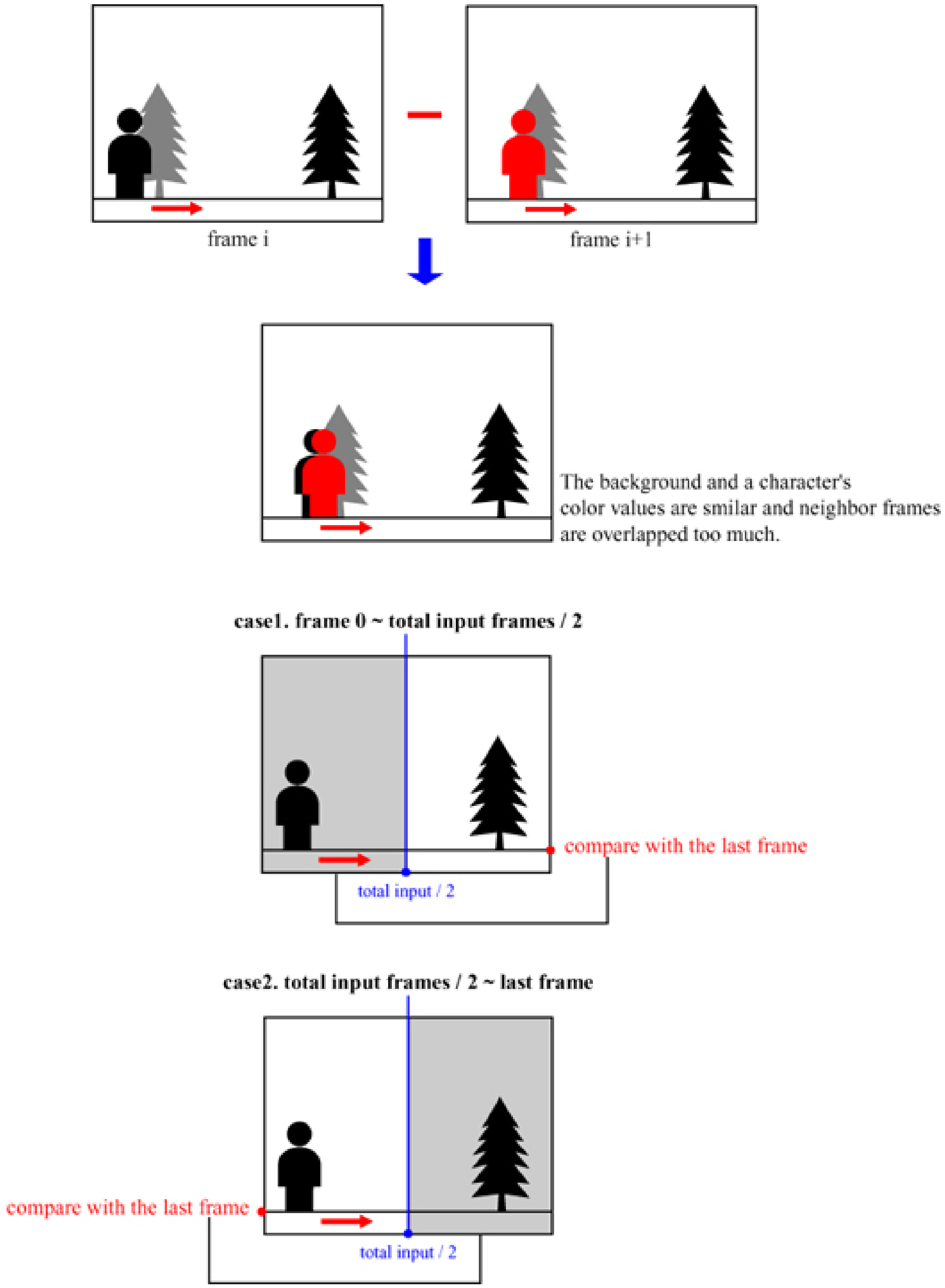


Fig. 42. Determination of the bounding box when the background is not the bluescreen.



The background and a character's color values are smilar and neighbor frames are overlapped too much.

Fig. 43. Extracting a character from the background.

# VITA

**Hye Mee Choi**

Visualization Laboratory
A216 Langford Center
3137 TAMU
Texas A&M University
College Station, TX 77843-3137

choihyemee@hotmail.com

## Education

| | |
|---|---|
| M.S. in Visualization Sciences | Texas A&M University, May 2004 |
| M.F.A. in Advertising Design | Hongik University, Korea, August 1996 |
| B.S. in Horticultural Science | Korean Sahmyook University, Korea, February 1993 |

## Research Interests

Lighting and Texturing
Animation
Layout

## Employment

| | |
|---|---|
| Graduate Assistant | Texas A&M University, September 2002 - May 2003 |
| Independent Contractor | Texas A&M University, March 2002 - May 2002 |
| Graphic Designer | Samsung Corporation, Korea, January 1997 - May 1998 |
| Graphic Designer | Hann Communications, Korea, 1995 |

## Honors

Chicago International Animation Film Festival, 2003
Downstream International Film Festival, 2003
IG(Independent Georgia) Music, Art, and Film Festival, 2003
Microcinema Independent Exposure, 2003
Industrial Light & Magic Inc./ Texas Aggie Alumni Award, 2001
Chosun Daily Newspaper Advertising Award, 1993-1994
Cheil Communications Inc. Students Advertising Award, 1993-1994