

# Categorical Organization and Machine Perception of Oscillatory Motion Patterns

James W. Davis

MS, Massachusetts Institute of Technology, 1996  
BS, University of Central Florida, 1994

Submitted to the Program in Media Arts and Sciences, School of  
Architecture and Planning, in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Media Arts and Sciences  
at the Massachusetts Institute of Technology

June 2000

© 2000 Massachusetts Institute of Technology. All rights reserved.

KOTCH

Author \_\_\_\_\_ Program in Media Arts and Sciences

Certified by \_\_\_\_\_ Aaron Bobick  
Associate Professor, Dept. Computer Science  
Georgia Institute of Technology  
Thesis Supervisor

Certified by \_\_\_\_\_ Whitman Richards  
Professor, Dept. Brain and Cognitive Science  
Massachusetts Institute of Technology  
Thesis Supervisor

Accepted by \_\_\_\_\_ Stephen A. Benton  
Chair, Departmental Committee on Graduate Students  
Program in Media Arts and Sciences

MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY

JUN 21 2000

LIBRARIES



# **Categorical Organization and Machine Perception of Oscillatory Motion Patterns**

**James William Davis**

Submitted to the Program in Media Arts and Sciences, School of Architecture  
and Planning on April 28<sup>th</sup>, 2000, in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy

## **Abstract**

Many animal behaviors consist of using special patterns of motion for communication, with certain types of movements appearing widely across animal species. Oscillatory motions in particular are quite prevalent, where many of these repetitive movements can be characterized by a simple sinusoidal model with very specific and limited parameter values. We develop a computational model of categorical perception of these motion patterns based on their inherent structural regularity. The model proposes the initial construction of a hierarchical ordering of the model parameters to partition them into sub-categorical specializations. This organization is then used to specify the types and layout of localized computations required for the corresponding visual recognition system. The goal here is to do away with ad hoc motion recognition methods of computer vision, and instead exploit the underlying structural description for a motion category as a motivating mechanism for recognition. We implement this framework and present an analysis of the approach with synthetic and real oscillatory motions, and demonstrate its applicability within an interactive artificial life environment. With this categorical foundation for the description and recognition of related motions, we gain insight into the basis and development of a machine vision system designed to recognize these patterns.

## **Thesis Supervisors**

Aaron Bobick  
Associate Professor, Dept. Computer Science  
Georgia Institute of Technology

Whitman Richards  
Professor, Dept. Brain and Cognitive Science  
Massachusetts Institute of Technology





## Thesis Committee

Thesis Reader \_\_\_\_\_

**Bruce Blumberg**

Assistant Professor, Media Laboratory  
Massachusetts Institute of Technology

Thesis Reader \_\_\_\_\_

**James Gould**

Professor, Dept. Ecology and Evolutionary Biology  
Princeton University

Thesis Reader \_\_\_\_\_

**Demetri Terzopoulos**

Professor, Dept. Computer Science  
University of Toronto



# Table of Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>INTRODUCTION.....</b>   | <b>9</b>  |
| 1.1      | CLAIMS .....   | 12        |
| 1.2      | ALTERNATIVES TO A CATEGORICAL APPROACH.....                      | 13        |
| 1.3      | REPRESENTATION AND RECOGNITION OF CATEGORICAL OSCILLATIONS ..... | 15        |
| 1.4      | SUMMARY OF CONTRIBUTIONS .....                                   | 17        |
| 1.5      | THESIS OUTLINE .....   | 17        |
| <b>2</b> | <b>MOTION RECOGNITION .....</b>                                  | <b>20</b> |
| 2.1      | ANIMAL MOTIONS .....   | 20        |
| 2.1.1    | <i>Motion discrimination of prey</i> .....                       | 20        |
| 2.1.2    | <i>Deception and motion camouflage</i> .....                     | 21        |
| 2.1.3    | <i>Conspecific visual communication</i> .....                    | 22        |
| 2.1.4    | <i>Miscellaneous activities</i> .....                            | 23        |
| 2.2      | HUMAN PERCEPTION: A DEMONSTRATION .....                          | 24        |
| 2.3      | SECTION SUMMARY .....  | 28        |
| <b>3</b> | <b>CATEGORICAL ORGANIZATION OF OSCILLATORY MOTIONS .....</b>     | <b>29</b> |
| 3.1      | MOTION SET .....   | 30        |
| 3.2      | A SIMPLE GENERATIVE MODEL.....                                   | 33        |
| 3.2.1    | <i>Structural parameters</i> .....                               | 34        |
| 3.2.2    | <i>Stylistic parameters</i> .....                                | 35        |
| 3.3      | COMPLEXITY ORDERING.....   | 36        |
| 3.3.1    | <i>Simple vs. complex motions</i> .....                          | 36        |
| 3.3.2    | <i>Sub-categories</i> .....                                      | 40        |
| 3.4      | BIOLOGICAL COMPARISONS.....                                      | 42        |
| 3.4.1    | <i>Simple as ubiquitous</i> .....                                | 42        |
| 3.4.2    | <i>Dance language of bees</i> .....                              | 46        |
| 3.5      | SECTION SUMMARY .....  | 48        |
| <b>4</b> | <b>COMPUTATIONAL IMPLEMENTATION .....</b>                        | <b>49</b> |
| 4.1      | RELATED WORK.....  | 50        |
| 4.2      | DATA REPRESENTATION .....  | 52        |
| 4.2.1    | <i>Point-light display</i> .....                                 | 53        |
| 4.2.2    | <i>Experiments</i> .....   | 54        |
| 4.3      | METHODS.....   | 61        |

|          |   |            |
|----------|---|------------|
| 4.3.1    | <i>Lowpass filtering</i> .....                        | 61         |
| 4.3.2    | <i>Pattern segmentation</i> .....                     | 63         |
| 4.3.3    | <i>Parameter estimation</i> .....                     | 64         |
| 4.3.4    | <i>Computational sub-category levels</i> .....        | 64         |
| 4.3.5    | <i>Periodicity calculation</i> .....                  | 66         |
| 4.3.6    | <i>Recognition methods</i> .....                      | 67         |
| 4.4      | SECTION SUMMARY .....                                 | 68         |
| <b>5</b> | <b>EXPERIMENTAL RESULTS</b> .....                     | <b>69</b>  |
| 5.1      | PARAMETER ESTIMATION .....                            | 69         |
| 5.1.1    | <i>Number of data samples</i> .....                   | 69         |
| 5.1.2    | <i>Gaussian noise</i> .....                           | 70         |
| 5.1.3    | <i>Non-idealized frequency and phase</i> .....        | 71         |
| 5.2      | MODEL CONFUSION .....                                 | 74         |
| 5.3      | GOAL-DIRECTED RECOGNITION .....                       | 78         |
| 5.4      | REAL VIDEO.....                                       | 85         |
| 5.5      | SECTION SUMMARY .....                                 | 90         |
| <b>6</b> | <b>INTERACTIVE ARTIFICIAL LIFE</b> .....              | <b>91</b>  |
| 6.1      | ARTIFICIAL LIFE ENVIRONMENTS .....                    | 91         |
| 6.2      | ARTIFICIAL HUMMINGBIRDS.....                          | 94         |
| 6.2.1    | <i>Virtual world</i> .....                            | 94         |
| 6.2.2    | <i>Implementation</i> .....                           | 99         |
| 6.2.3    | <i>Perception and action</i> .....                    | 101        |
| 6.3      | SECTION SUMMARY .....                                 | 101        |
| <b>7</b> | <b>CONCLUSION AND SUMMARY</b> .....                   | <b>103</b> |
| 7.2      | SIGNIFICANCE AND CONTRIBUTIONS .....                  | 105        |
| 7.3      | FURTHER QUESTIONS AND PROBLEMS WITH THE APPROACH..... | 107        |
| 7.4      | FUTURE DIRECTIONS.....                                | 108        |
| <b>8</b> | <b>ACKNOWLEDGEMENTS</b> .....                         | <b>112</b> |
|          | <b>APPENDIX A</b> .....                               | <b>113</b> |
|          | <b>APPENDIX B</b> .....                               | <b>120</b> |
|          | <b>APPENDIX C</b> .....                               | <b>125</b> |
|          | <b>BIBLIOGRAPHY</b> .....                             | <b>126</b> |

# 1 Introduction

We live in a dynamic world full of motion. To interact effectively (as man, animal, or machine), it is necessary to distinguish and identify the intentional or meaningful movements of animate creatures from other non-relevant motions in the environment (e.g. falling leaves). The perception of various *types* of movement in the world is important for attaining knowledge of the surrounding environment and for survival and reproduction [19, 106]. Observers actively exploring their world use motion to give them information on both the structures in the world and about their own movement [19]. Motion is a very primitive cue to detect prey, as shown with frogs that have specific motion detectors to identify particular types of insects [59]. To survive these attacks from predators, prey must be able to detect and perform evasive maneuvers, possibly using motion cues as well. The identification of patterns of motion is also used for the communication of intention. Male spiders use stylistic motions to differentiate themselves to a female as a potential mate rather than a potential meal [43]. This is but one of many examples of the role of motion in guiding more social behavior. Most animals use information about the *activities* of others in courtship, mating, aggression, parental care, and play [19]. Here the animal must be able to perceive what other animals are *doing* and adjust its own behavior accordingly.

Human vision seems to be very much more general, although it clearly contains a variety of special-purpose mechanisms [69]. Early in life, a newborn infant responds to the movement of objects dangled in front of it, but appears to be unaware of the same object at rest [106]. More involved motion perception includes a dance instructor, gymnastics coach, or air-traffic controller who must be able to interpret the higher-level movements in complex dynamic visual information. The common and stereotypical head, hand, and body gestures of people reflect the more social perception of human motion. The ways in which people interact with others involve many further complexities arising from culture and language, but human social interaction requires the same basic abilities to recognize and detect what other people are doing [19]. We can further illustrate this notion with the classic film by Heider and Simmel [49] that shows three simple geometric shapes moving and interacting with each other. Viewers of this film report a rich story-like interpretation only from the actions of the moving shapes. Similar interpretive abilities can be seen with the immediate perception of human motion from dynamic point-light displays [55] or from simply watching a children's puppet show. For this to happen, there must be regularities in the movement patterns to allow such strong interpretations of the motions (possibly distinguishing between animate and in-animate motions). To be able to discern these types of

*special* motion patterns in an unambiguous manner, the movements must have physical regularities to which the visual system is tuned.

Any one animal does not exhibit the full possible range of movement behaviors capable of its body design. There are only so many ways an animal can walk, run, or jump. This is particularly true for communicative acts, for if it performed an infinite variety of motions to communicate, then the perceiving animal would not be able to “get the message” or “figure it out”. Darwin himself even noticed this commonality of behavioral expression across a large variety of animals and man [29]. This therefore suggests that *categories* of limited motion exist and that they each have regularities to which perception systems may be tuned. Perhaps it is the *non-categorical* behavior exhibited of prey that makes it especially attractive and “stand out” to predators, as when lions more likely attack the gazelle that appears injured or aged as revealed by its abnormal gait pattern.

To offer just a few of these natural motion categories, we mention three broad classes of motion: impulse, stochastic, and oscillatory. Impulse motions include step-force movements like a jump or a jab. Stochastic motions make up those movements such as vibrating or shaking. Oscillatory motions in their most ideal form are best described as smoothly repetitive sinusoidal motions. We can easily find these types of motions in a variety of behaviors (communicative and non-communicative) across the animal kingdom. In the first categories for example, animals aggressively dart at enemies, jump between branches, and zig-zag away from predators. Vibratory shaking movements frequently occur when an animal bathes, when a baby bird begs for food, or when an animal trembles from fear. Oscillatory motions are common and can be seen when birds bob and sway their head while walking, when body limbs swing or flap for walking or flying, and when a person gestures back-and-forth with the head and hands. We present additional examples in the Appendices of these types of movements as taken from bird behavior. During some communicative movements, repetition of irregular motions may be used as an overarching mechanism to generate visual signals. Other low-level categories of visual motion have also been proposed and defined based on boundaries within movements (starts, stops, force discontinuities) [85], rather than on the explicit pattern itself.

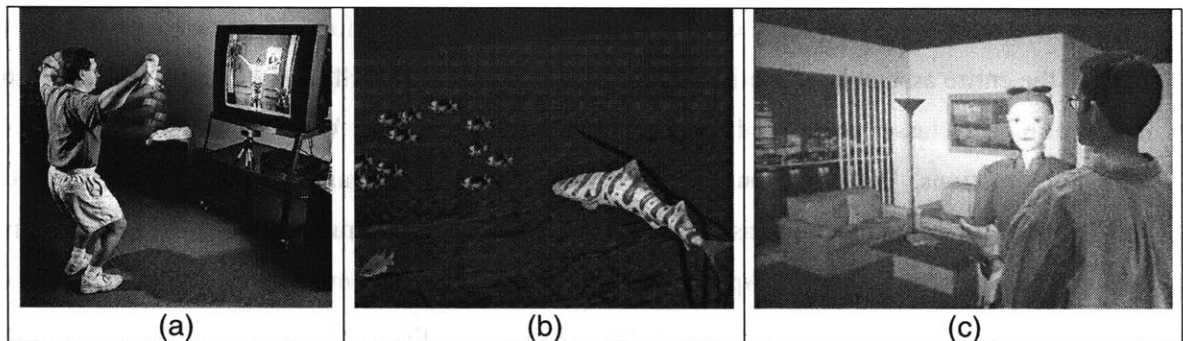
As mentioned previously, we see these motion categories appearing in several different types of behavior. With this behavioral non-exclusivity, the resulting perceptual representations for the categories are applicable to a more general perceptual framework as needed for daily activities

and survival. Upon specialization of these general categorical models, particularly tuned versions can be created for more specific and directed communication tasks (e.g. courtship). The fact that many of the communicative signaling motions of animals are actually derived from more basic non-communicative behaviors [100] strengthens the notion that the recognition systems for communicative motions are likely to be derived from the corresponding basic perception systems. Here the perceptual mechanism specific to certain communicative motions leverages off the more general system based upon the needs for the basic non-communicative behaviors and perceptions. Thus a representation based on the categorical nature of motion is most parsimonious and supports the larger tasks necessary of a more general perception machine. Though it is believed that animals do not all use the same representations for general vision [69], we may however find a commonality in the representation *goals* if several species respond to the same (or similar) category or instance of motion. The actual biological implementation mechanism perhaps is different, but the nature of the visual information required for recognition by the various species remains the same – they all must solve the same visual problem.

Clearly our initial categorization of motion is very crude. Finer sub-categories can be documented, creating a very large catalog for types of animate motions. Hence we cannot hope to analyze the entire assembly of motions in this thesis. Instead we will limit our scope to one of the most prominent classes associated with communicative behavior. We have chosen the class of oscillatory motions. We begin our focussed study with an examination of oscillatory motion patterns used for communication as described in the behavioral literature for *several hundred* bird species (See partial listing in Appendices A-C). It was quite evident that certain types of oscillatory motions were used by many species as a form of visual communication. These patterns include, for example, simple up-and-down bobbing, circular patterns, and looping flights. We believe that such distinctive motions are based on more cognitive patterns [68], rather than governed strictly by motor capability, due to the large number of species that perform many of these motions (from insects to spiders to birds to man) and the anatomical diversity of performance (using the head, body, arm, foot, tail).

The research here strives to construct a meaningful categorical description of oscillatory motion that can be exploited for the development of an artificial perception system that can identify these common motion patterns. Consider the formidable challenge of designing and implementing a perceptual virtual world (as in [22, 31, 98], See Figure 1) where users and artificial creatures freely communicate and reciprocate intentions. It is natural to inquire how one might design a

perception system to recognize meaningful movements to enable artificial creatures to watch and categorize these patterns (of a person, animal, or artificial creature). The methodology derived here for a single motion category is general in that the *approach* is not specific to the characterization of only one particular category. Our approach is to model a set of categorical motions appearing across species by noticing that regularities exist for the related (categorical) motions which can be used to help organize a perceptual model for recognition of the patterns. The categorical recognition model for the target set of oscillatory motions is based upon the recovery and matching of the structural parameter regularities from a simple characterizing sinusoidal model. The motions are hierarchically ordered in complexity using the notion of increasing specialization in the limited sinusoidal model parameters. The resulting organization is partitioned into sub-categories highlighting the prominent forms of oscillatory specialization. These ordered sub-categories are used as the basis for the specification and layout of computation for the perceptual recognition mechanism. We can then examine the capabilities of this approach with controlled experiments and by using the models within an interactive virtual world.



**Figure 1.** Example systems where artificial creatures require perception for interaction with other artificial creatures or human beings. (a) Davis' virtual aerobics trainer. (b) Terzopoulos' artificial fish. (c) Cassell's Rhea.

## 1.1 Claims

We postulate that a single underlying foundation describing a set of related patterns motivates the design of a perception system that uses the shared structural features in the motions for categorical recognition. We assert three claims related to this manner of categorical motion perception:



- Claim 1**      Very specific patterns of motion are used by animals as a means of visual communication.
- Claim 2**      Many of these motion patterns are oscillatory and can be characterized with a simple (categorical) sinusoid model.
- Claim 3**      A hierarchical organization of the sinusoidal motions can be used to partition and specify the types of computation for recognizing the class of motion patterns.

With Claim 1, we state that various types of animal species use specialized motion patterns for the task of visual communication with other species. We have already briefly mentioned examples of this and will present further evidence from ethological studies supporting the claim by showing uses of motion for identification, deception, and communication by predators, prey, and conspecifics. Interestingly, certain types of patterns appear across species. In Claim 2, we state that many of the patterns in Claim 1 are oscillatory and can be modeled with simple sinusoids. As a result of examining the behavior literature of several hundred bird species, we discovered and compiled a set of *oscillatory* motions found in bird behavior (in other animals as well). We will present a simple sinusoidal model with very special and limited parameter values that characterizes these motions. Lastly in Claim 3, we propose that a sub-categorical organization for the motions can be used as a means to specify the generalized form of computation for a categorical recognition system. We will present a computational implementation that is organized based on the motion sub-categories and show the results using this approach.

## **1.2 *Alternatives to a categorical approach***

Instead of addressing the *categorical* nature of the motions, we could alternatively build an individual and specific recognition model for each motion on an “as needed” basis. If each pattern to recognize were of a completely different nature, this approach may be the only course of action. But, as the number of patterns of interest grows, a corresponding increase in the set of distinct models (and processing) occurs. Due to clear similarities between so many motions and the limits of the types of these patterns, it is unreasonable to think that perception is based on a different

model (or approach) for each motion. Designing such specialized systems is undesirable because the result is to have only a set of mimicry devices with no understanding of the underlying processes [69]. For example, template matching one pattern to another is an extremely context sensitive and ad hoc approach, looking for correspondences at a point-wise level with no underlying model parameters to exploit. Here any minor change to the input (e.g. scale) can have drastic results in the matching result.

Another approach is analysis and recognition of the motions within generic high-dimensional parameter spaces (e.g. principal components). In such methods, exhaustive training examples for the entire *class* are collected and decomposed into a set of basis motions capturing the appearance of the data. But these approaches do not necessarily recover the important structural foundation in the data. In the case of oscillatory motions, linear decomposition methods using the sinusoidal trajectories (position traces) as primitives do not capture the true underlying functional parameters for they are inherently non-linear (i.e. frequency cannot be linearly decoupled from the basic sinusoidal position generator). Another training-based matching method commonly used for recognition is a Hidden Markov Model (HMM). Though less sensitive to deviations in the signal than with straight trajectory correspondence (template matching), the method is generally used for matching an input trajectory to a trained state-based model from examples of a particular “class” of motion. The term class is used here in a more specific sense of the word, meaning a collection of perturbed versions from an ideal example. It does not relate to multiple motions that may be qualitatively different in appearance but similar in structural nature. Therefore the HMM does not encode the broad categorical generalization needed to capture a broad range of structurally related, but visually different, motions. Thus, one needs a new HMM for each motion pattern to recognize. Again, these methods described design specialized systems that result in having only a set of mimicry devices with no understanding of the underlying processes and constraints in the motions.

Fortunately, there exist structure and regularity in many motions (See [85] for a related description of qualitative changes and boundaries of motion). We believe that the success of categorical perception relies on model representations that reflect these structures and regularities of the world that transfer to the visual image (similar to ecological optics of [45] and natural modes of [82]). Categorical representations may also enable the prediction other non-observable properties from these observable regularities [13]. With the notion of a category, the structural constraints for the class of motion are made explicit, and model generation for various categorical

motions is made simple by choosing from these fixed parameter values and constraints permitted within the category. In other words, it's the “structural relations” that define the class of movements, not the statistics of multiple training examples (as with principal components and HMMs). Thus creating a new pattern model from a category requires no training, unlike methods that demand multiple new examples for each additional model. A categorical change may therefore be considered a form of one-shot learning [38].

This categorical approach is also less sensitive to slight deviations in the signal (e.g. from viewing condition or performance) for it has a notion of parameterized structure that can be used to aid recognition. For example, if the categorical method is concerned with the structural foundation of frequency and phase parameter *relationships* for sinusoidal trajectories, then the analysis can be lenient on the remaining parameters in the signal (e.g. amplitude, speed). Thus recognition can be tolerant of particular deviations in the signal. We see this in actuality with the existence of the structural form (e.g. a figure-8) and stylistic performance (e.g. fast or slow) of the motions. The system should be able to separate and identify the qualitative pattern without influences from the stylistic performance. Training-based approaches (even those that are parameterized) using the raw trajectory information, on the other hand, will have difficulty when they encounter a new slightly-deviated version of the expected pattern if it is not accounted for during training (e.g. if a new style is encountered). Thus to be effective for recognition, the training class would need to encompass all possible structural and stylistic properties that are capable of being produced in the generation of the patterns. The exploitation of these strong regularities and parameter constraints within the perceptual window enhances the competency for the model [81]. As a result, the structural category approach is more descriptive (and compelling) as a means for the representation and perception of related motions.

### **1.3 Representation and recognition of categorical oscillations**

The categorical approach begins with a compilation of related patterns that a perception system is to identify. In this research, we focus on the common oscillatory motions as highlighted in bird behavior. Once such a motion set is chosen, a parameterized structural model is required (though it may be challenging for a mathematical description of some other categories). This offers a unified description for multiple movements. For the resulting organization to be useful for perception, the model parameters and regularities for the motions need to be features that can be

visually extracted. For the oscillations, we chose three parameterized sinusoidal functions with parameters of amplitude, frequency, phase, and translation. With a parameterized characterization for the motion set, a hierarchical organization on the model can be imposed to show the increasing specialization in the motions as ordered from simple to complex. Because we have a single sinusoidal model for the motions, we can identify the more complicated motions as being composed of the simpler patterns with additional parameters and/or constraints (e.g. up-and-down + horizontal translation = undulation). Thus we can place the patterns in a hierarchy with the simplest motions at the top (with the least number of active model parameters) and the most complex motions at the bottom (most parameters). The downward branching within the hierarchy reflects those motions that are most closely related. Since there exist multiple levels and parameter/constraint commonalities (regularities) within the hierarchy, we can cluster the motions into groups that exhibit the prominent specializations appearing in the motion class. This partitioning corresponds to a sub-categorization of the motion set.

The computational significance of this categorical framework is that a single model hierarchically organized into sub-categorical complexities can be used to specify the type and organization of computation needed to recognize the motions. This sharing of computation (assigned to the sub-categories) is more desirable than redefining computation specifically for each pattern to recognize. The result of using such a computational layout is that only basic computations are required for simple motions, and additional computations are used only for the components found in the more complex patterns. Thus, the approach shares computational routines where appropriate. The implementation for the oscillatory motions first extracts individual patterns cycles in the input position trajectories using sign-changes in velocity that indicate boundary locations. Then Fourier analysis is used to estimate the sinusoidal parameters from the segmented trajectories, identifying only the parameters that are required by the model (sub-category) being tested. After extracting the required sinusoidal parameters from the input trajectories, their values are matched to the relations and constraints defined in the category description. The computations for extracting the pattern, estimating the parameters, and matching to the models are all categorically organized and implemented.

We stress here the importance of understanding the nature of categorical motions more than on the actual implementation details. There are no claims that the implementation mimics actual biological mechanisms used for recognizing the motions. Rather the goal here is to demonstrate instead a simple recognition algorithm that describes and captures the categorical patterns using

the constraints and organization inherent to the motions. The nature of the computations underlying the task depends more upon the computational problems that have to be solved than upon the particular hardware mechanism in which the solutions are implemented. In other words, the implementation details don't matter so much provided that they do the right thing [69]. In our case, how those localized computations are in fact implemented is not as important as how they are organized.

### **1.4 *Summary of contributions***

The major contributions of this thesis are:

- The identification and compilation of related oscillatory motions that are used for communication throughout the animal kingdom.
- A rank ordering of the oscillatory motions derived from a descriptive model that has a correlation to an ordering of bird motions.
- A method to specify the types and layout of computation for recognition from the underlying categorical organization in the motions.
- An experimental result showing that trajectories of oscillatory patterns of human arm movements are best represented with position rather than velocity data.
- A real-time computer implementation that recognizes the oscillatory patterns based on a structural categorical organization without the need for statistical training on large sets of examples.

### **1.5 *Thesis outline***

We have motivated and introduced a method for recognizing categorical motions. We presented three claims to our approach. First, that special types of motion patterns are used by animals for

visual communication. Second, that many of these patterns are oscillatory and appear sinusoidal. Third, that a hierarchical organization of the motions can be used to specify the types of shared computation needed to recognize the patterns. We also discussed the advantage of a categorical approach over other training-based methods. The remainder of this thesis is devoted to presenting the categorical model and the results using the approach with various forms of oscillatory motion. That this approach can be utilized for categorical recognition of the motion patterns will be demonstrated.

We begin in Section 2 by further examining the ethological roles of motion recognition for animals with respect to locating prey, deceiving prey and predators, and communicating with conspecifics. Additionally in Section 2.2, we present evidence for the ability of human perception to recognize motion patterns in the absence of strong image features, emphasizing the importance of the motion pattern for recognition. A related computational vision method motivated by this ability is also briefly described.

In Section 3, we derive a categorical organization for an important class of oscillatory motion. We begin by identifying a set of oscillatory motions used throughout the animal kingdom for communication (especially by birds). A simple sinusoidal model with very specific and limited parameter values is then presented in Section 3.2 that characterizes all of the oscillations, where the motions can be decomposed into structural and stylistic components. In Section 3.3, a hierarchical ordering on the motions is derived within the category, ranking the patterns from simple to complex based on the number of model parameters and constraints for the motions. Additionally, the category is partitioned into sub-categories that reflect the prominent forms of oscillatory structural specialization. In Section 3.4, we compare this ordered categorical structure with relations of animal motions to seek a biological correlation to the model organization.

Section 4 presents the corresponding perceptual model implementation for recognizing the motions from their organized categorical structure. In Section 4.1 we review related work in computer vision associated with recognizing periodic and oscillatory motions. In Section 4.2, we outline experiments with human arm motion to suggest the form of the underlying data representation for the motion signal. An infrared tracking apparatus is described in Section 4.2.1 for creating point-light motion displays from user motions for these experiments. The remaining details of the categorical implementation methods are described in Section 4.3.

Results of the categorical implementation are presented in Section 5. Evaluation of the approach involves analysis of the models (and associated issues) using the contexts of synthetic data, user-generated point-light displays, and real video. Experiments and results focussing on parameter estimation are first presented (5.1). Due to the differing complexities in the recognition models from the category structure, we next examine the nature of the model confusions for each of the motions (5.2). Lastly the models are tested with real video examples of oscillatory motion (5.4). The combination of these tests is used to quantify and assess the strength of the framework.

In Section 6, we further explore and demonstrate the categorical models with an interactive artificial life environment. We show that the oscillatory perception system outlined in this research is useful in constructing autonomous virtual creatures capable of varied interactions connecting movement patterns in the real world to creature behaviors in a virtual world. Previous artificial life and interactive systems that motivated the demonstration are first described in Section 6.1. We then present in Section 6.2 an artificial hummingbird world designed particularly for this research as a means of using the recognition models within an entertaining, interactive virtual environment. The design (6.2.1) and implementation (0) of the virtual world are described.

In Section 7, we conclude with a summary of the significance, contributions, and future directions of this work, emphasizing the methodology proposed for constructing machine-based vision systems concerned with recognizing meaningful categories of motion. Collections of avian motions used for this research to compile the motion set and for comparison of the categorical organizations are listed in detail after the summary in Appendices A, B, and C.

## 2 Motion Recognition

Sensory experience is made up primarily not of things, but of dynamic forces [3], and thus humans and animals detect real-world structure by detecting changes along the physical dimension [42]. A dog or cat may be resting peacefully, unimpressed by the splendor of colors and shapes surrounding it, but as soon as anything stirs, its eyes turn to the spot and follow the course of motion; young kittens seem completely at the mercy of moving things as though their eyes were tied to the objects in motion; as for humans, the effectiveness of mobile advertising need only be mentioned to show the effectiveness of motion [2]. Thus it seems reasonable for beings to additionally be able to discern “special” patterns of movement to better interact within the environment. It is clear that there exists a strong coupling between perception and action, where particular movement patterns are unambiguously recognized and acted upon by animals and humans. From an ethological standpoint, the benefit of being able to perform and recognize certain patterns of motion is clear [7, 15, 43, 46]. Consider the roles of motion used by animals for locating prey, deceiving predators, and communicating with conspecifics.

### 2.1 *Animal motions*

#### 2.1.1 *Motion discrimination of prey*

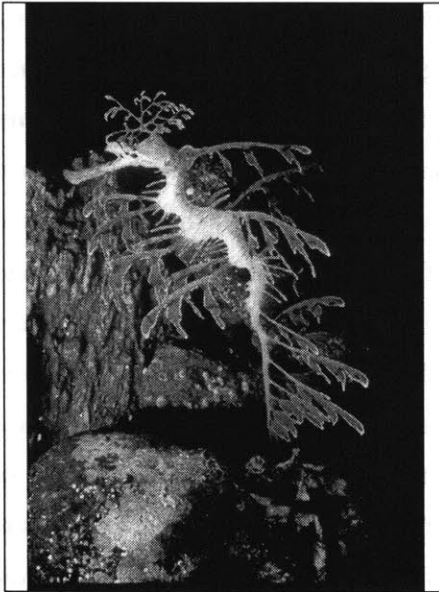
Carnivorous animals must identify, pursue, and capture prey for survival. For identifying prey, some predators simply respond to low-level motion cues. This appears in reptiles, which don't see color that well, that mostly respond to movement. For example, snakes have very limited powers of sight and a moving object is necessary to arouse their attention [20]. The frog also responds to motion [59]. Some predators can discriminate an easier prey to catch by recognizing abnormal movement patterns of an individual within a group. For example, lions can pick out the individual gazelle whose gait pattern is abnormal (reflecting injury, sickness, or age), thus offering an easier meal to catch [84]. Perhaps the abnormal motion patterns do not belong to the expected category of motion for the animal.



### 2.1.2 Deception and motion camouflage

Another effective method of hunting is to use lures to attract the unsuspecting prey. Perhaps the best-known example of using lures is given by the anglerfish. This creature has a long, stiff, thread-like lure with a fleshy tip that it wiggles in front of its mouth. Small fish are attracted by this quivering lure, and the anglerfish then opens its huge mouth and sucks in the unsuspecting fish with the in-rushing water [65]. Similarly, the alligator snapping turtle has a fleshy pink, worm-like lure on the floor of its mouth that it wiggles to attract fish [65]. Conversely, there are those prey that cleverly perform particular movements that are used to distract a predator. For example some birds (e.g. killdeer, plover) distract a predator from their young by moving away from the nest in a hurt fashion with a wing fluttering as if it were broken or damaged. This “broken-wing display” draws the attention of the predator away from the nest towards a prey that is believed to be an easy catch [6]. Only when the predator is sufficiently far from the nest does the bird discontinue its persuasive acting. Africa's two-headed snake deceives predators using its head-like tail as a distractor. Its real head has no distinguishing features, whereas its tail has the coloring and shape of a typical snake's head. As the snake moves away from a confrontation, the tail mimics the defensive strikes made by the head of an actual retreating snake [76].

As with animals camouflaged in background colors or textures, sometimes it may only be the motion of the animal which surrenders its guise. But, there are animals that are camouflaged in their movement as well as appearance. For example, the Leafy Sea Dragon [47] looks and moves like a piece of seaweed carelessly floating through the water (See Figure 2). This creature is of interest here mainly because its movement patterns mimic that of a soft plant slowly drifting and bending through the water, rather than any noticeable animal-like motion. We consider this type of movement a form of “motion camouflage”, where the animal no longer moves as a living creature (with *active* motion) but rather mimics that of an inanimate object (in this case, the *passive* motion of a plant floating through the water). Such a form of movement camouflage, in addition to its leafy, plant-like appearance, has clear advantages when considering predation on/by such an animal. Snakes similarly mimic inanimate motion by moving with a swaying locomotion to appear like the background motion of bushes swaying in the wind [40].



**Figure 2.** Pattern and motion camouflage of the Leafy Sea Dragon.  
(Picture courtesy of the New England Aquarium.)

### 2.1.3 *Conspecific visual communication*

Animals also use motion to recognize and communicate with others of their own species (conspecifics). In the simplest case, it has been shown that young chicks respond more to the downwards pecking motion of a cutout model of the mother rather than to her pose [104]. Another example is the important role of the motion pattern of the male Grayling butterfly for finding a female counterpart [101]. Also, *Sceloporus* lizards each tend to bob their head in a specific rhythm for identification of their own species [70], and various species of fiddler crab similarly have different claw waving tempos [26]. The female spider's instinct is to seize any moving animal small enough for her to taste and to kill. If a male spider would simply walk up to the female, she would attack and kill him relying on her feeding instinct. It is therefore essential for male spiders to have some way of identifying themselves to females. This is accomplished by the male performing specific movement patterns to the female to announce his presence as not another meal [20, 43].

Many animals perform movement “displays” as a form of visual communication. Very often, displays serve to attract attention by moving/waving colored and textured body parts [7]. Elaborate courtship displays are performed by many animals to arouse the mating instinct [43]. A

few well-known bird courtship displays include the Blue-footed booby “marching” in-place (alternately lifting his bright-blue feet up-and-down) [6] and the Ruby-throated hummingbird flying in a pendular U-shuttle path [96]. Courtship and threat displays are many times the same or similar for a given species. Thus a single display is able to court a female and ward off rival neighbors simultaneously [7]. There are also essential similarities in territorial defense rituals and alarm signals for many species [43]. Thus many communicative display behaviors (e.g. courtship, territory, alarm) can coincide to common signal types. In general, many display movements are ritualized or exaggerated forms of related or more basic behaviors (e.g. the swaying display motion of the guinea pig is reminiscent of its mating behavior [15]).

Much of this ability is not learned behavior, but rather innate to the animal itself and can be combined or re-shaped from a genetic standpoint. In normal courtship for instance, diving ducks like the redhead throw their head back toward the tail, while dabblers like the shoveler stand up in the water with their head down on the breast and grunt or whistle. A male hybrid from this combination was found to have both tendencies and tried to combine them when he reached the courtship stage. He rose in the water like the shoveler but threw his head back like the redhead [21]. Similarly, a male hybrid of a prairie chicken and a sharp-tailed grouse first boomed like a chicken and the other chickens approached. Then the hybrid raised his tail and danced like a grouse [21]. Both of these hybrids were not successful in attracting a mate.

#### *2.1.4 Miscellaneous activities*

Distinctive motion patterns can also be seen in the more daily activities of animals. For example, coots bob their head when swimming or walking [32]. The kingfisher similarly bobs its head when perched [44]. Woodpeckers are known for their stylistic undulating flight pattern [48]. Phalarope shorebirds spin around in circles on the water to bring up small organisms when feeding [60]. Some animals confined in zoos are seen pacing repetitively back-and-forth.

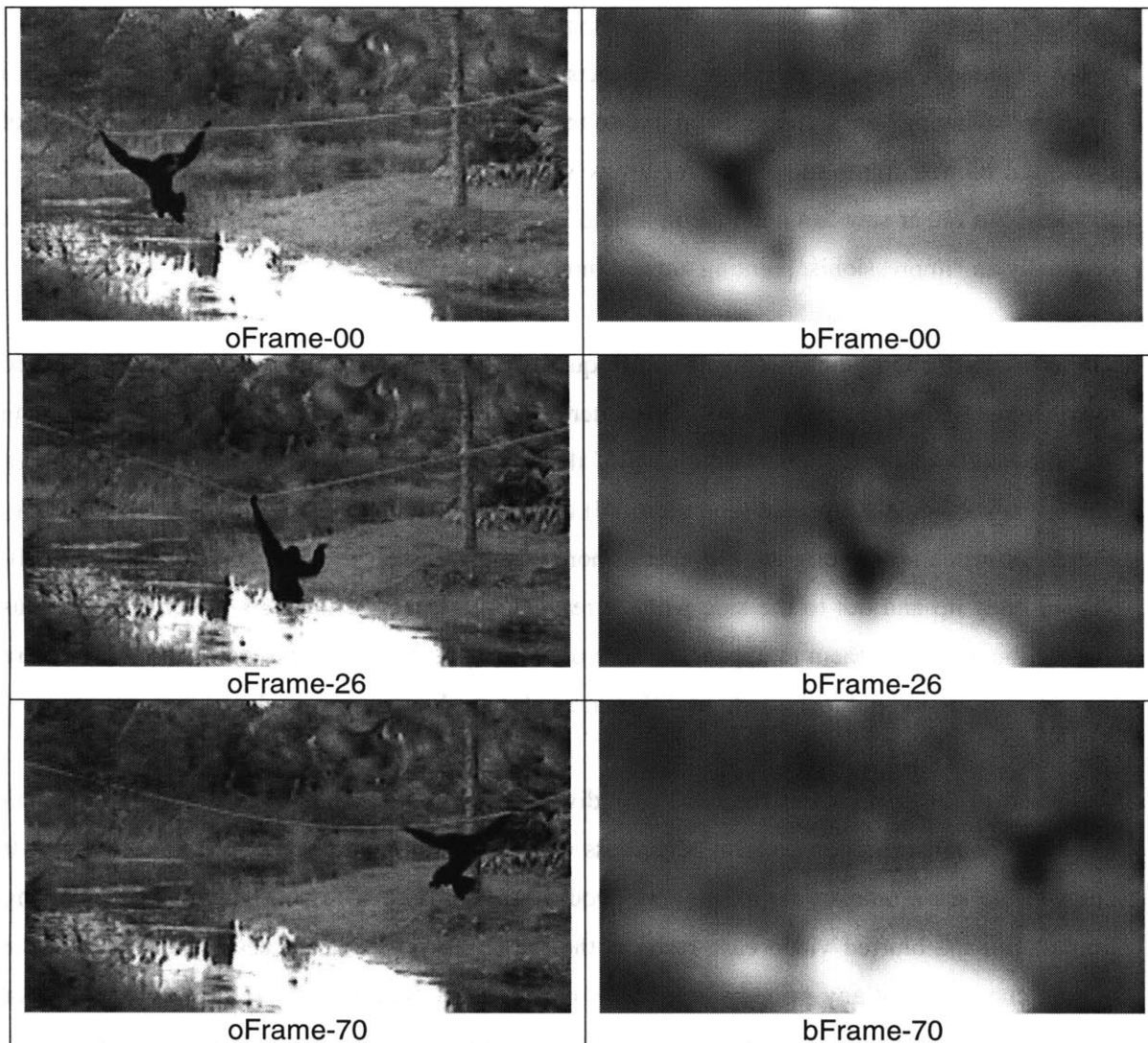
Animals also exhibit interesting motions during play. For example, chimpanzees [37], gorillas, and panda bears [9] cartwheel, looping across the ground when playing. Also, pronghorns run in circles [37], coyotes wave their head [8], and bees even perform brief looping flights near the hive entrance as form of play flight [46].

An interesting use for the common head swaying, bobbing, and thrusting in certain animals is to generate image motion for judging depth or distance [19]. For example, a locust makes a side-to-

side swaying of the head to generate image motion for predicting distances to jump between two surfaces. Similar head movements are seen in gerbils, which make vertical bobbing movements of the head before jumping over a gap. Also well known are the head bobbing and head thrusting forward-and-backward (relative to the body) of doves, pigeons, and chickens, which is believed to give them a measure of visual depth.

## ***2.2 Human perception: a demonstration***

As illustrated in watching a mime performer with only “show” and no “tell” [1], people can simply view and recognize the movements without surrounding environmental contexts or other multi-modal input. Certain motion patterns have been shown to be unambiguously recognized from their own regularity of movement. For example, people can easily identify a collection of brightly moving dots in a dark scene as a set of lights attached to the joints of a walking person [55]. In fact, babies have been found to stare longer at such point-light displays rather than if the dots moved randomly [41]. To further illustrate this motion recognition ability, we performed a simple experiment asking people to describe the motions they perceived in very low-resolution (blurred) sequences. The low-resolution and the original higher-resolution movie clips of motions were shown separately to two groups of people and then the viewers were asked to describe the dominant motion seen in each sequence. Five people were shown several short normal resolution movie clips of different motions, and five other people were shown the same clips but with the frames blurred to extreme as to obscure the detail and context of the images as much as possible. The set of movie clips included a flitting hummingbird, a hopping kangaroo, a shaking hand, a back-and-forth running bird, a jumping bird, a head-bobbing goose, and a swinging gibbon.



**Figure 3.** Key frames from the gibbon swinging sequence. The left and right columns show keyframes from the original and blurred sequences. When viewers were asked to identify the scene in the first blurred image (bFrame-00), no one was able to discern the gibbon. (Filmed by the author with permission of the Brevard Zoo, FL.)

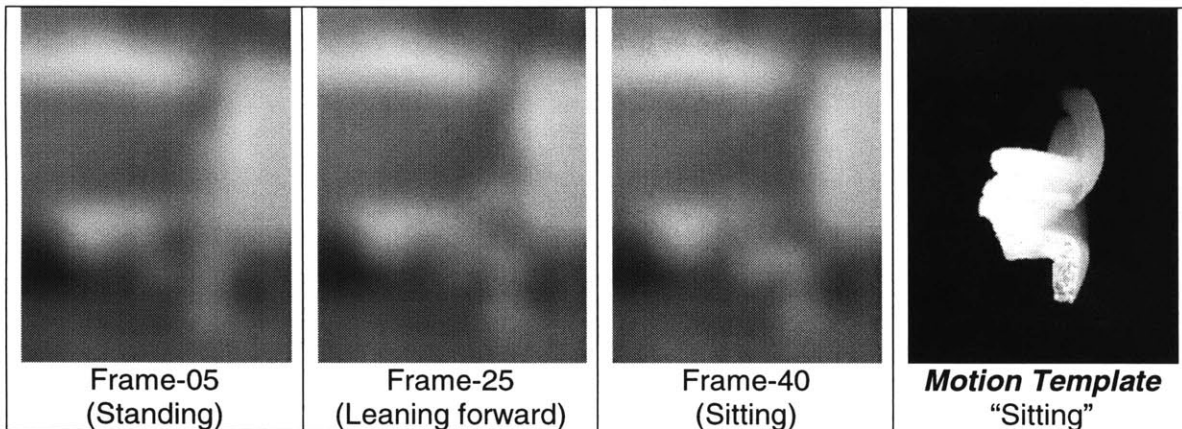
The left column of Figure 3 shows a few key frames from the original sequence of a gibbon “swinging” along a rope, and the right column of the figure shows the corresponding blurred frames. The viewers of this video were each asked simply to describe or label the dominant motion exhibited in the sequence. Most subjects viewing the original sequence described the motion using terms or phrases related to swinging (e.g. “swing(ing)”, “brachiation”, “W-shape, but with smooth bottoms”). Similarly, subjects viewing the blurred sequence reported seeing a monkey “swinging” or “jumping from tree to tree”, even though when describing the contents of the initial blurred frame (displayed for a few seconds before the animation) they did not see a monkey, but rather saw “an antelope in a field”, “maybe a plot of trees in the distance”, or most originally an “impressionist painting viewed through a dirty window”.

These results for the gibbon sequence were qualitatively similar to the other blurred/un-blurred movie clips shown to the viewers. One interesting difference arising from the two separate viewing conditions was found in the case of the side-to-side motion of the bird. In the blurred video, viewers mainly reported (and gestured) an animal moving repeatedly to the left and right. But those watching the original clip could more clearly see that the bird actually tilted its bell-shaped body up and out to the side when it reached the extent of the right or left motion, thus biasing the motion to them to appear as more of a swinging type of back-and-forth motion. Here the video resolution did make a slight difference in the qualitative motion description.

The result of this experiment showed that individuals share similar motion descriptions and can recognize common motion patterns regardless of surrounding contextual cues in the imagery. If one were to see a monkey on a vine, one would *expect* to see it “swing” during the following frames. But the blurred video case restricts the viewer to perceive the action mostly from the motion cues, without any prior knowledge or expectations. The above test shows that the prior context of a scene (or situation) is not necessarily required to recognize or label the motion.

We mention briefly here that this ability to recognize motion patterns from extremely low-resolution imagery was the basis for our earlier work on template-based motion recognition. In [30], we presented a real-time computer vision approach for representing and recognizing simple human movements. The motivation for the approach was based on how easily people can recognize common human movements (like sitting or push-ups) from low-resolution (blurred) imagery as shown in the above experiments (see a blurred sitting sequence in Figure 4), where there is an almost total lack of recognizable image features. Again, people have no problem

identifying the sequence when played as a movie, thus emphasizing the role of motion for recognition. Accordingly, the computational recognition method relied on using “patterns of motion” rather than particular image features (e.g. head, arm) as the representation for recognizing human motion. In that method, the space-time image volume containing the image motion is collapsed (compressed) into a single 2-D template that still perceptually captures the *essence* of the movement and its temporal structure (See motion template in Figure 4). For recognition, seven higher-order moments [52] are extracted from the template, and then statistically matched to stored model examples of different movements. Though the method is powerful for recognizing large spatially distributed motions (as demonstrated with a large database of movements), it does not consider the structural relations between similar movements (i.e. two similar motion patterns do not share the same model). In that method, each motion requires a different model for recognition (means and variances of the moment feature vectors). The categorical approach to be developed here seeks out the relations between motion patterns to provide a more general methodology for understanding the motions and the construction of recognition systems designed to identify these patterns.



**Figure 4.** Key-frames from a low-resolution sequence of a person sitting down as viewed from the side. The last image shows the motion template representation of the sitting movement (created from the higher-resolution image sequence).

### ***2.3 Section summary***

In this section we motivated the importance of motion recognition and have provided several examples of animal motions illustrating the roles of motion in nature throughout everyday experience. Such an ability to distinguish and identify various forms of animate motion is paramount in their existence. We also demonstrated the ability of human perception to recognize motion patterns in the absence of any contextual information, stressing the important role of motion in our perception. In the following sections, we present our computational model for representing and recognizing categorical oscillatory motions.



### 3 Categorical Organization of Oscillatory Motions

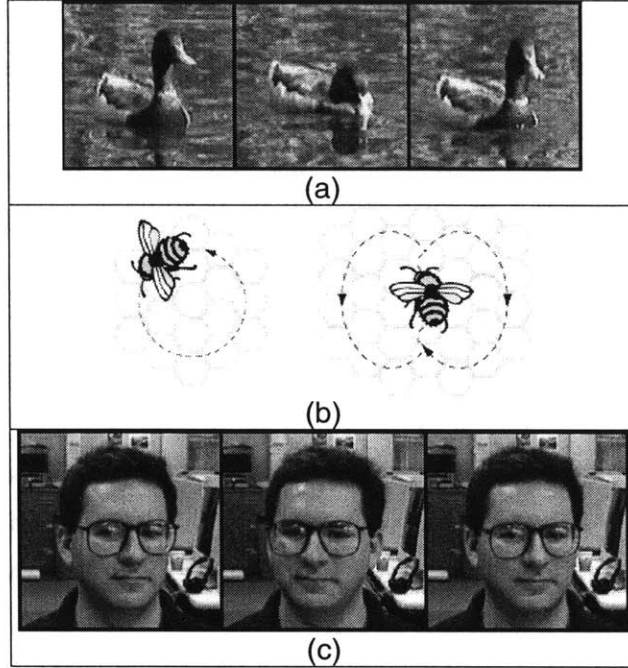
In this section, we first compile an important set of oscillatory motions that are used throughout the animal kingdom as a means of communication. We explore a simple sinusoidal model with characteristic regularities that captures the categorical nature of the oscillations with structural constraints and stylistic capabilities. We then show that there is an underlying ordering of these motions based on the choices and number of model parameters, where the increasing specialization in the model parameters is exploited to hierarchically rank the motions in complexity. The resulting organization is used to partition the motion category into sub-categories that highlight the prominent forms of oscillatory specialization. We then present a comparison of our structural ordering to observations of related animal movements. Our perceptual approach developed later for recognizing these patterns will reflect this structural categorization for the motions.

For this research, we performed an extensive searching and cataloging of several communicative oscillatory animal motions and behaviors (including several hundred bird species) in an attempt to find an important set of related motions found across species to demonstrate the model and implementation. We chose to examine *communicative* and *intention* display motions because they are particularly interesting in that they have goals associated with them and their consistency with respect to form, speed, and rhythm makes them favorable for characterization [74]. By the term *intention*, we refer to the ethological interpretation and definition, rather than the AI descriptions (with respect to beliefs, goals, and plans) where there is not a single definition that is widely accepted. Ethologists refer to intention movements as the preparatory action (e.g. motion, posture, color-flashing) which always precede a subsequent behavior (e.g. attacking, fleeing, copulating), but may be performed without completing the following behavior. Such movements are derived from specific behavioral actions and provide information to the receiver about the actions the sender would like the receiver to believe it will do next [100]. These preparatory movements generally become some form of display signal and thus act as a communicative mechanism. Thus we are interested in perceiving single communicative action units that correspond to some current (or immediate future) behavioral state, rather than attempting to understand the internal beliefs and longer term goals of the creature.

Studying the basic mechanisms of specialized movement employed throughout nature makes the characterization of motion more easily distinguishable from random patterns. Since many animal displays are derived from non-display behaviors [100], studying these communicative display patterns may be useful in understanding the more general structure of movement patterns and behaviors. Therefore, it is worthwhile to examine communicative motions, for it may also reveal shared features of the perceptual and mental models used for representing and perceiving these movements in everyday experience.

### **3.1 Motion set**

As previously described, many behaviors throughout the animal kingdom consist of performing and recognizing specialized patterns of motion. Birds seem especially well endowed with communicative display movements that are visually rich and diverse. We examined the communicative motions for several different species and it was obvious that certain types of oscillatory motion were pervasive throughout bird behavior (across several bird species), where a common and predictable feature of their display movements is rhythmic (oscillatory) repetition. For example, Mallard ducks (*Anas platyphynchos*) bob their head *up-and-down* to a female [94] (See Figure 5.a), Ruby-throated hummingbirds (*Archilochus colubris*) swing in a pendular *U-shaped path* [96], and Peregrine falcons (*Falco peregrinus*) *loop* through the air [96]. Outside the avian domain, we also see similar oscillatory movements used for communication. For instance, spiders wave their palps *up-and-down* as a form of courtship [25], honey bees dance in a *circle* or *figure-8* to signal the presence and location of a food source to the hive [108] (See Figure 5.b), Sceloporus lizards identify one another from *push-ups* [70], and chimpanzees sway from *side-to-side* as a form of threat (or for courtship) [36]. Even humans *nod* or *wag* their head as a simple responsive gesture (See Figure 5.c). Oscillations are quite common in biological systems and can also be seen in animal locomotion [24, 75]. It is interesting that certain types of oscillatory motion patterns appear across such a wide variety of species.



**Figure 5.** Communicative oscillatory motions in nature. (a) Mallard duck head bobbing. (Filmed by the author at the Boston Public Gardens.) (b) Honey bee dancing in a circle and figure-8. (c) Person slightly waving his head ‘No’.

An important issue with regards to cataloging this type of motion is to define what exactly constitutes an oscillatory motion pattern. We define “oscillation” as swinging *steadily* back-and-forth with a definable period in each spatial dimension utilized for movement. We refer to more general “periodic” motion as generic movements happening at regular intervals, with no smooth continuity enforced (i.e. not steady). For instance, oscillatory motions can be described using sets of simple sinusoid functions, but this is not so for all periodic motions. Oscillatory motions are thus a subset of the larger class of periodic (repetitive) motion and the sinusoidal description cannot extend to this entire class. We speculate that general periodic motion models could contain two sub-models, one for the individual motions and one for their repetitive linking. For example, one can repetitively draw a square with four linear motions (up, right, down, left) that have a periodic continuation (by drawing the box multiple times without lifting the pencil). According to our definition, this repetition is periodic, but not oscillatory and cannot be described with simple sinusoids. We note that though general periodic motion can be found in communicative motions, we commonly find oscillatory patterns for communication. Also, a simple sinusoidal model makes it easy for characterization of the motions. We therefore focus on this particular subset of periodic motion.

When we examine the literature seeking oscillatory movement behaviors for sinusoidal characterization, we need to be careful to accept those motions best described by our oscillatory definition. For example, we can find birds that “flick” their tail *side-to-side* (e.g. Downy woodpecker [95]). We do not consider this motion a pure oscillatory pattern. Though repetitive, this motion appears to be a combination of a horizontal oscillation and an impulse-like flick motion. Thus a motion such as this can not be fully characterized by sinusoids (may be represented with a two-state coupled mechanism, repeating a left-flick, right-flick, and so forth). Similarly, we can see cross-cultural eyebrow flashing motion of people where the eyebrows are raised, held for some duration, and then quickly lowered [34]. This motion may be compactly described by “moving the eyebrows up-and-down”, but this is more of a triphasic eyebrow flash (due to the raising, holding duration, and lowering of the eyebrows). Again, this would not be considered oscillatory with a sinusoidal model.

To address this oscillatory selection problem, we attempted to gather those motion descriptions that expressed an oscillatory “smoothness” using descriptions such as “rhythmic” with terms “swaying”, “bobbing”, and “wagging”. Repetitive motions described with “flicking”, “jerking”, and “snapping” were not included. There will nonetheless be some concern for certain terms with semantic ambiguities (e.g. does moving back-and-forth mean going left-right or forward-backward?). Obviously, some of the particular motions catalogued may in actuality be more periodic rather than oscillatory, but an attempt was made to try and minimize this error. Another related issue is whether or not the perceiver of these display patterns has a percept of oscillatory motion or in fact some other response. In a recent personal experience, a bird was encountered that kept flying out from a tree branch in a small circular path with another bird of the same species in proximity (clearly a display). Certainly the *generative* motion pattern was a circular oscillation (one cycle), but halfway through the circle motion the bird flashed a bright white patch under its tail. Such color flashing is common in the animal kingdom as a display signal. Most likely the *perceiving* bird was responding to the tail flashing rather than to the strict circular motion pattern. Therefore we must be aware of the possibility that some displays may not be perceived as movements, but instead may have a particular component associated with them (e.g. a flash) that acts as the true releaser.

In Table 1, we list a set of those oscillatory motion patterns which appear widely throughout the animal kingdom as a form of communication and that in particular are used frequently by birds as

a form of display behavior. These motions are regularly reported for birds in common field guides (e.g. [17, 32, 33, 64, 92, 94-96]) with phrases similar to “rapid head bobbing”, “circles over territory”, “slow spiral upward”, “wags tail up-and-down”, “repeatedly flies in tight figure-8 pattern”. Example bird species that perform these motions as displays are additionally listed in the table along with the motions. The *up-and-down* and *side-to-side* motions are simple 1-D “bobbing” and “swaying” motions, as in head nodding or wagging. The *circle* pattern is a circular 2-D motion around a center point, as exemplified by a male tern that walks round and round a female during courtship. The *spiral* motion is a 3-D curve moving up or down about a central axis, also referred to as a “helix”. We use the term *spiral*, as this is the most common description given in the behavioral literature (e.g. “spiraling up into the air”). The *loop* pattern is a “prolate cycloid” curve, appearing like somersaults. The *undulate* motion is wave-like motion (a sine wave). The *figure-8* pattern is typically a sideways figure-8 movement, also referred to as a “lemniscate”. The U-shuttle motion repetitively traces out a bowl-shaped “parabola” opening upward with finite extension.

| Table 1. Example oscillatory avian displays* |   |
|--|---|
| Motion                                       | Birds   |
| <b>Up-and-down</b>                           | Bluejay [94], Pileated woodpecker [32], Dark-eyed Junco [96].           |
| <b>Side-to-side</b>                          | Snowy owl [17], Northern fulmar [32], Red-eyed vireo [94].              |
| <b>Circle</b>                                | Gray phalarope [33], Sparrowhawk [33], Woodlark [33].                   |
| <b>Spiral</b>                                | Mourning dove [95], American woodcock [96], Isabelline wheatear [33].   |
| <b>Loop</b>                                  | Booted eagle [33], Mockingbird [94], Redpoll [33].                      |
| <b>Undulate</b>                              | Northern goshawk [96], Yellow-bellied sapsucker [32], Desert lark [33]. |
| <b>Figure-8</b>                              | Piping plover [32], Peregrine falcon [96], Black-eared wheatear [33].   |
| <b>U-shuttle</b>                             | Ruby-throated hummingbird [96], Hobby [92], Northern harrier [64].      |

\* Three example bird species are given for each motion, though the motion may be performed by a wide variety of species.

### 3.2 A simple generative model

The movements in Table 1 are all comprised of simple oscillatory motion, and can be modeled as sinusoids. This basis seems especially appropriate here, as most of these patterns are reminiscent

of Lissajous figures and appear sinusoidal. All the oscillatory motions in our set can be described by three simple parameterized sine-wave generators of the form:

$$\begin{aligned} x(t) &= A_x \sin(2\pi f_x t + \phi_x) + B_x t \\ y(t) &= A_y \sin(2\pi f_y t + \phi_y) + B_y t \\ z(t) &= A_z \sin(2\pi f_z t + \phi_z) + B_z t \end{aligned} \quad (1)$$

where  $(x,y,z)$  is the 3-D Cartesian location of some point feature over time (perhaps the body in motion or a color patch on a moving body part),  $A$  is the amplitude of the wave,  $f$  is the frequency,  $\phi$  is the phase shift, and  $B$  is the translation. The above description simply dictates the shape or path of the motion over time (i.e. the motion pattern). Obviously more complex and dynamical models could be used, as used to characterize locomotion [24] and handwriting [51] (see [81] for additional examples of natural computation models), but they too must obey the underlying oscillatory nature as given by the sinusoidal model.

### 3.2.1 Structural parameters

Table 2 shows the parameter settings for this sinusoidal model needed to characterize our set of oscillatory motions (idealized in their purest, planar form). Looking closely at the parameters in the table, we find very specific and limited values. For these motions, the only 2-D frequency ratios are  $\{1:1, 1:2\}$  and the only relative phases (locking  $\phi_x=0$  for reference) are  $\{0, \pm\pi/2, \pi\}$ . In particular, the phase-difference relation for circular motions (circle, spiral, loop) must obey  $\phi_x - \phi_y = \pm\pi/2$ , for figure-8 the relation must be  $2\phi_x - \phi_y = \{0, \pi\}$ , and for U-shuttle the constraint is  $2\phi_x - \phi_y = \pi/2$ .

For looping, its translation  $B_x$  is constrained by the product of its corresponding amplitude and frequency ( $2\pi f \alpha_x$ ), otherwise swinging occurs. This boundary constraint for looping can be shown using derivative analysis. For rightward looping to occur, the motion derivative in the horizontal x-axis must be periodically negative (to loop back). Thus the bounds of  $B_x$  can be found by determining when the derivative is always positive (generating a swinging pattern instead of looping). Setting the derivative of  $x(t)$  to be greater than or equal to zero ( $\alpha_x 2\pi f \cos(2\pi f t + \phi_x) + \beta_x \geq 0$ ) corresponds to  $\beta_x \geq 2\pi f \alpha_x$ . Thus  $B_x$  for looping must be in the range  $[0 \dots 2\pi f \alpha_x]$ . For translation in the opposite direction, we have  $[0 \dots -2\pi f \alpha_x]$ . We

don't have "hard" boundary for the other translation motions of spiral and undulate, although an extreme translation for these motions will distort the overall visual shape of pattern. Therefore we only list a non-zero translation ( $\neq 0$ ) in Table 2 for spiral and undulate.

| <b>Table 2. Oscillatory Motions and Model Parameters</b> |            |            |            |           |       |       |          |            |            |                                 |          |       |
|--|------------|------------|------------|-----------|-------|-------|----------|------------|------------|---------------------------------|----------|-------|
| Motion   | Amplitude  |            |            | Frequency |       |       | Phase    |            |            | Translation                     |          |       |
|  | $A_x$      | $A_y$      | $A_z$      | $f_x$     | $f_y$ | $f_z$ | $\phi_x$ | $\phi_y$   | $\phi_z$   | $B_x$                           | $B_y$    | $B_z$ |
| Up-and-down  | 0          | $\alpha_y$ | 0          | –         | $f$   | –     | –        | 0          | –          | 0                               | 0        | 0     |
| Side-to-side   | $\alpha_x$ | 0          | 0          | $f$       | –     | –     | 0        | –          | –          | 0                               | 0        | 0     |
| Circle   | $\alpha_x$ | 0          | $\alpha_z$ | $f$       | –     | $f$   | 0        | –          | $\pm\pi/2$ | 0                               | 0        | 0     |
| Spiral   | $\alpha_x$ | 0          | $\alpha_z$ | $f$       | –     | $f$   | 0        | –          | $\pm\pi/2$ | 0                               | $\neq 0$ | 0     |
| Loop   | $\alpha_x$ | $\alpha_y$ | 0          | $f$       | $f$   | –     | 0        | $\pm\pi/2$ | –          | $[0 \dots \pm 2\pi f \alpha_x)$ | 0        | 0     |
| Undulate   | 0          | $\alpha_y$ | 0          | –         | $f$   | –     | –        | 0          | –          | $\neq 0$                        | 0        | 0     |
| Figure-8   | $\alpha_x$ | $\alpha_y$ | 0          | $f/2$     | $f$   | –     | 0        | $0, \pi$   | –          | 0                               | 0        | 0     |
| U-shuttle  | $\alpha_x$ | $\alpha_y$ | 0          | $f/2$     | $f$   | –     | 0        | $-\pi/2$   | –          | 0                               | 0        | 0     |

Model parameters are shown for the oscillations generated by sinusoidal functions ( $X(t) = A \sin(2\pi f t + \phi) + B t$ ). Values  $\alpha$  and  $f$  correspond to variable amplitude and frequency values, respectively. Slots with – are non-applicable parameters due to corresponding zero amplitudes.

Although many other distinctive values could exist for the sinusoid parameters (e.g.  $\pi/3$ ,  $4f$ , etc.), they are not seen in these oscillatory motions. The parameter values shown in the table uniquely define the motions, and yield the simple oscillatory patterns. Thus the use of other parameters, even just slightly different from those listed, cause a more complex repetition pattern to form over time. Such special values for this model suggest that generative or structural regularities [63, 82, 109] underlie this class of movement. Though some other motion categories may not be as easily modeled mathematically for their structural form, most will likely have similar modes in the patterns that may include regularities in time (speed), position, or even performing limb.

### 3.2.2 Stylistic parameters

In addition to the structural parameter relations described above for the qualitative pattern, there exist parameter values designed for specific stylistic variation in their performance. In the motions, stylistic information can be encoded into the amplitude and/or frequency parameters

resulting in selective and immediate recognition from the same structural pattern (e.g. as used for species identity [70, 88, 97]). The translatory component in the motions may also be stylistic (e.g. looking for tight looping vs. spread-out loops), and thus the traveling amount may be selectively wired into the perception system. For example of an amplitude stylization, consider the pendular U-shuttle movements for the Black-chinned hummingbird (*Archilochus alexandri*), Broad-tailed hummingbird (*Selasphorus platycercus*), and Calliope hummingbird (*Stellula calliope*). They all perform the U-shuttle movement, but each with differing heights ranging from about 4.5-7.5, 6-12, to 18-27.5 meters, respectively [97]. Given that these species reside and overlap in the same West Coast territory, this difference may act as an isolation mechanism [99] during interactions between the species. These motion patterns are qualitatively similar (i.e. pendular swinging:  $f_x = \frac{1}{2} f_y$  and  $2\phi_x - \phi_y = \frac{\pi}{2}$ ), but differ in the amplitude of the height (i.e.  $\alpha_y$  from shallow to steep). We can therefore specify the oscillatory motions in terms of the sinusoidal model with descriptions of the constrained *form* (e.g. the U-shuttle) and a selective *style* (e.g. the height amplitude).

### 3.3 Complexity ordering

Given the set of categorical motions and the sinusoidal descriptive model, the question remains whether we can organize the motions based on their relations in the model representation. An ordering of the motions is necessary if we wish to understand how the motions are related to each other within the category, making clear the developments of more complex patterns from simpler motions. Although many orderings on the motions may be possible depending on the specific criteria used for the complexity of the patterns, we present here a hierarchical ordering for the motions based on increasing specialization (complexity) in the number of sinusoidal model parameters and constraints.

#### 3.3.1 Simple vs. complex motions

Given the model representation (sinusoids) and the structural constraints (specific frequency ratios, phases differences, and translations) for the oscillatory patterns, we can now further organize the motions by their amount of increasing specialization exhibited in the model parameters. In terms of motion complexity, we focus on the structuring of the generative model parameters rather than the performance medium (e.g. in the air, on the ground, the performing limb). These additional factors though may suggest complementary organizations for the motions (e.g. motions in the air may be more complex to perform than on the ground). For our ordering



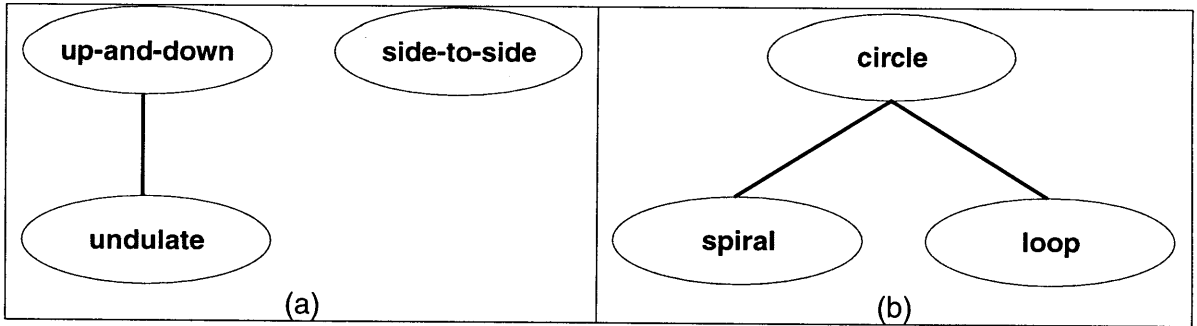
based on model parameter specialization, we make explicit the number of parameters and constraints used to generate the patterns within the category. A hierarchical organization is formed by noticing that some motions are combinations of more primitive movements or have additional parameters/constraints.

We choose a *hierarchical* ordering for the motions rather than a strict listing or table of the model parameters. Though a listing of model parameters for the motions, as shown in Table 2, is a complete specification of the patterns, the table format itself is a tedious representation that is confusing to understand with no obvious visible trends or relations. The main use of a table representation would be as a simple look-up device (i.e. index the model parameter choices by the motion label). This ignores the categorical relations between the motions. Instead, a hierarchical organization for the motions is a simple and intuitive representation making clear how the motions are related and organized. Since the complexity organization for the motions is based on the number of active model parameters and constraints (i.e. more parameters/constraints = more complex motion), we can define a hierarchy for the motions with the depth in the tree reflecting the complexity from the number of model parameters and constraints in the motions. Thus, as we move down through the hierarchy, we add more parameters and constraints to the motions. This one-dimensional complexity relation is simple to understand. Also, the branching within the hierarchy reveals those motions most closely related to one another. Such hierarchical organizations have similarly been used for interpreting or explaining visual scenes [54, 67, 83].

As defined in [83], a type of hierarchical organization referred to as a structure lattice takes context-sensitive, primitive concepts about structural regularities, and composes them to produce a set of possible configuration states (e.g. spatial layouts, motion patterns). The lattice displays a partial ordering of the categorical states, where the ordering is derived by noticing that some states are special or limiting cases of others. The lattice shows the specialization relations (i.e. non-accidental features [66, 83]) between the structures in the world that appear in the representation given the observable properties. At the top of the partial ordering is a state of random generalization. This is the least structured node. Then at each level below, a dimension of specialization is incrementally imposed until reaching the bottom node(s) in the ordering. The bottom node(s) is completely specified with regularities (most structured). We follow this approach using the structural sinusoid parameters to order our oscillatory motions making explicit the relations and complexities of the motions. At the top of the ordering are the least constrained (or least structured) motions (i.e. the simplest patterns). At each level below, a dimension of

specialization on the parameter structure (e.g. adding more parameters) is incrementally imposed on the simpler patterns to create the more complex motions. The bottom leaf nodes in the hierarchy thus present those motions most completely specified with regularities (most structured and complex).

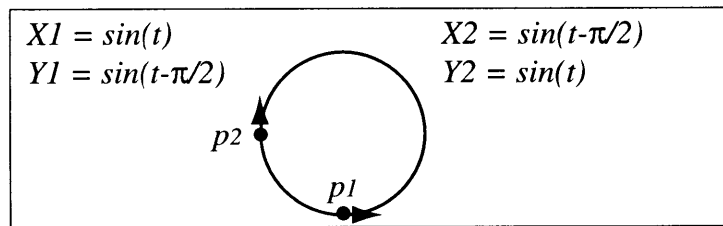
The least constrained movements in our category according to the sinusoidal model are clearly the up-and-down and side-to-side movements, due to their 1-D nature and small number of *active* model parameters (default non-active parameters values are zero). If the remaining zero-value parameters for the 1-D motions (zero amplitude and translation) were non-zero, this would enforce additional constraints with the other active parameters (i.e. relative phase and frequency constraints) making the pattern more complex overall. Because we have two states at the top of the ordering, we refer to the organization as a hierarchy rather than incorrectly as a structure lattice (requiring a single root node). The sinusoidal model also states that adding a translation to up-and-down (up-and-down + translation) generates the more specialized undulation motion (See Figure 6.a). Similarly, the more complex spiraling and looping patterns are derived from the simpler circle motion (See Figure 6.b) by adding a translation (2-D circle + translation), regardless of the direction of movement (horizontal, vertical). But the relationships for circle, figure-8, and U-shuttle are not as simply identified from the model, though clearly they are derived by combining the simpler up-and-down and side-to-side motions (horizontal motion + vertical motion = 2-D motion) with particular constraints.



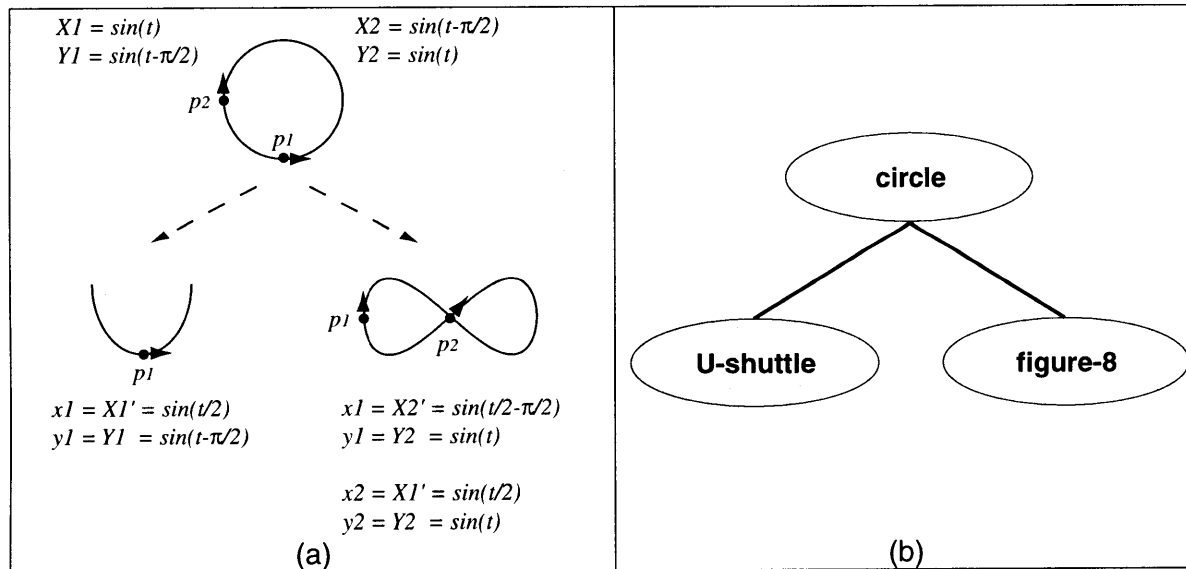
**Figure 6.** Derivations of motion patterns from simpler forms. (a) Undulate is derived directly from up-and-down motion. (b) Spiraling and looping are derived from circle motion.

The 2-D circular motion can be generated by combining the top-level up-and-down and side-to-side motions with a phase difference constraint of  $\phi_x - \phi_y = \pm\pi/2$ . The two possible signs of the phase difference correspond to clockwise and counter-clockwise directions (bi-directionality). We

know that both directions are possible for certain animal species, as exemplified by the male bridled tern that circles the female, periodically reversing direction [17], and by the honey bee's round dance where it runs in circles reversing direction often after one or two circles [108]. Thus two sinusoidal generators (one set for each direction) with a single phase value ( $-\pi/2$ ) flipping between the x and y functions can be utilized to account for the bi-directional circle motion (See Figure 7). We can then create a parsimonious organization for the circle, figure-8, and U-shuttle motions noticing that U-shuttle and both directions of the figure-8 can be easily derived from the bi-directional circle functions. We simply need to recombine the individual x and y circle functions with a frequency-doubling constraint (doubling the frequency in one of the two dimensions,  $f_x = \frac{1}{2}f_y$ ) to generate the U-shuttle and both directions for the figure-8 pattern (See Figure 8).

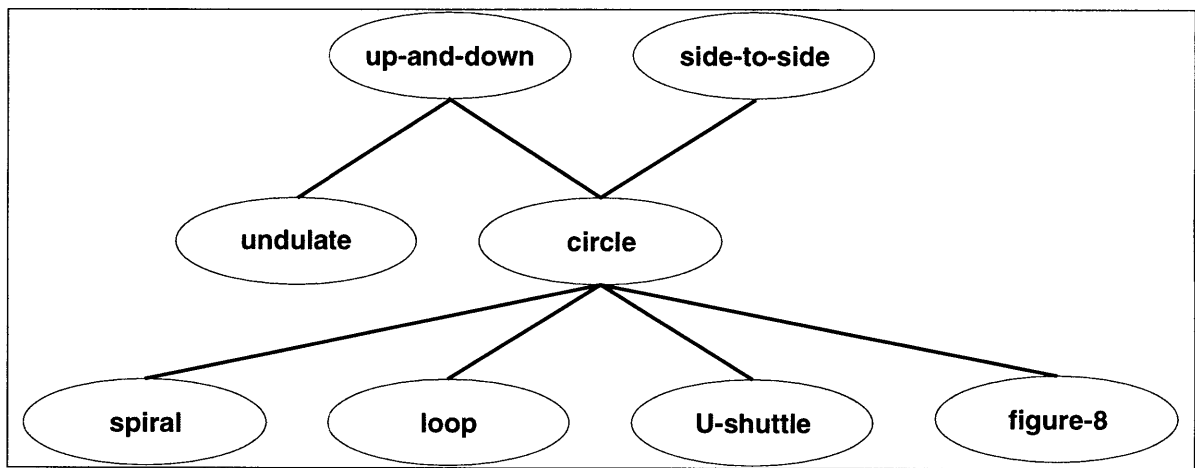


**Figure 7.** Bi-directional circle motion with two sets of function generators.



**Figure 8.** Ordering for circle, U-shuttle, and figure-8 motions. (a) The generating functions present for a bi-directional circular motion are recombined with a frequency-doubling constraint to generate both U-shuttle and figure-8. (b) The ordering on the motions resulting from the specialization in (a).

The resulting hierarchical ordering for the entire class based on the ranking of parameters in the oscillatory motions is shown in Figure 9. The ordering consists of three basic levels, with simple up-and-down and side-to-side motions at the top of the hierarchy, undulate and circle at the middle level, and the spiral, loop, U-shuttle, and figure-8 at the bottom of the ordering. The general evolution of the patterns begins with simple 1-D motions, which then combine to give circular motion followed by diversification into the translatory circles (spiral, loop) and frequency-doubling patterns (U-shuttle, figure-8). The undulate motion is somewhat unique with its 1-D (simple) and translation (complex) components.

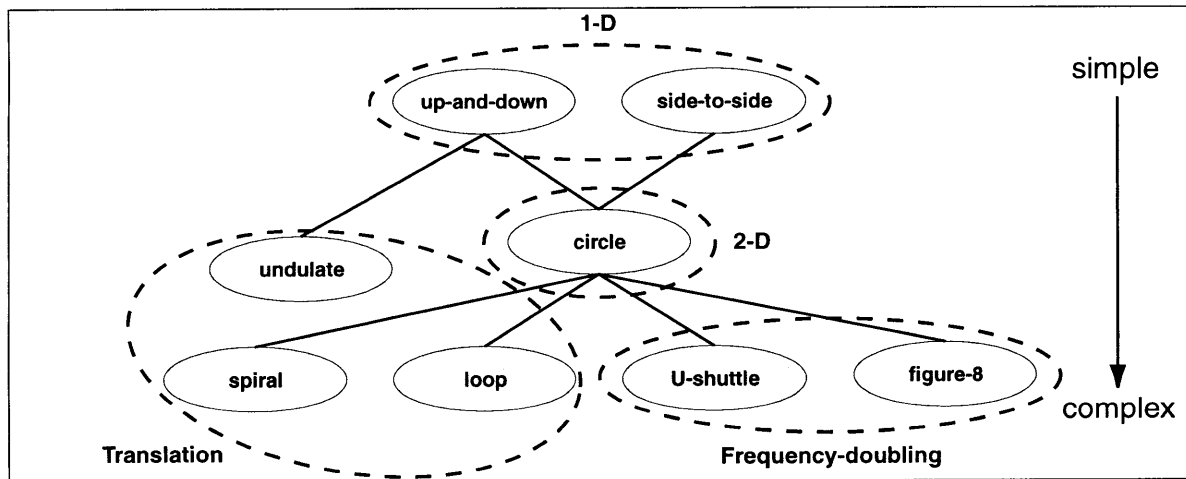


**Figure 9.** Complete hierarchical complexity ordering for the motion set.

### 3.3.2 Sub-categories

With this ordering, we can derive sub-categories that reflect the prominent types of specialization in the patterns. Noticing a common translation component, we can group together undulate, spiral, and loop into a *translation* sub-category. Though undulate is not at the same hierarchy level as spiral and loop, we feel the presence of a complex translation is significant in its sub-categorization. The later recognition mechanism based on these sub-categories will have a component to deal with this translation specialization, and therefore undulate (with a translation) should belong to this same sub-category. With a common constraint of a 2-D frequency ratio of 1:2, U-shuttle and figure-8 can be grouped into a *frequency-doubling* sub-category. This leaves circle motion to represent the earliest introduction of *2-D motion*. And lastly, the up-and-down and side-to-side movements correspond to a simple *1-D motion* sub-category.

Thus, the hierarchical structure for the category orders the individual motions and the sub-categories from simple to complex in relation to their structural specializations (See Figure 10). Using this hierarchical mapping, the relations between the motions and sub-categories become clear, with the depth in the hierarchy reflecting the complexity of the motions (number of active parameters). According to this hierarchy, for example, up-and-down is simpler than circle motion, which in turn is simpler than U-shuttle. Similarly, we can say that 1-D motion is simpler than 2-D motion, which is simpler than frequency-doubling motion. As we mentioned earlier, the undulate pattern resides on the second level in the hierarchy, but we group it in the translation sub-category with the third level translatory patterns. This collectively makes undulate more complex than the other second level circle motion, though individually undulate *should* be simpler than, say, looping. This sub-categorical organization for the motions will be used as a template to specify the types and layout of computation in the recognition model, where the sub-categories for the class are cast into localized computations ordered in complexity.



**Figure 10.** Hierarchical complexity ordering showing sub-categorization of the prominent specializations. The hierarchy orders the motions and sub-categories from simple to complex.

### **3.4 Biological comparisons**

In this section we compare the resulting categorical motion organization described in the previous section to observations and relations of animal movements. The purpose here is to seek biological organizations of the motions that support the theoretical relations derived from the model specialization. Looking to observations of animal motions offers an objective organization constructed completely independent of our research goals with which to compare our motion ordering. First, we examine the assertion that the simplest movements are the most ubiquitous across species (with particular emphasis on birds) and then present a similar partial ordering for certain motions as observed in the communicative dance language of honey bees. These results help to support our original intuitions from the hierarchical organization.

#### **3.4.1 Simple as ubiquitous**

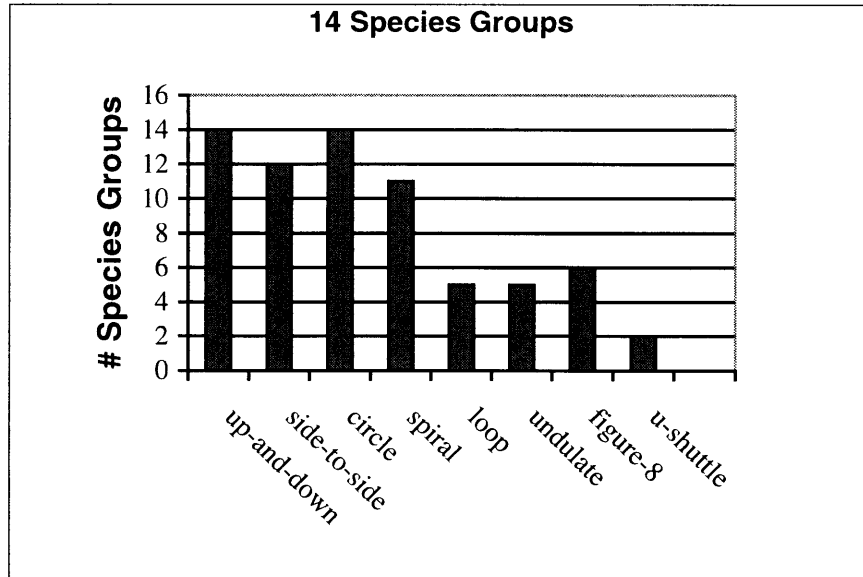
We expect that the simpler top-level motions in the hierarchy as constructed from the model (i.e. up-and-down, side-to-side), should indeed be the most generic given that these motions are likely to require the least amount of cognitive and motor control. Hence, they should be the most universal across various animal species. This assertion is supported by the ubiquity of these simple motions as used for communication across a large variety of animal species, including insects, spiders, lizards, birds, and chimpanzees. For example, during courtship, Salticid spiders wave their palps up-and-down [25], adders raise up and sway their head from side-to-side [20], giant water bugs pump their body up-and-down [93], fiddler crabs move from side-to-side and wave their big claw up-and-down [26], and Sceloporus lizards perform push-ups [70]. Other communicative motion displays include turtles swaying their head side-to-side [105], woodpeckers bobbing and swaying [32], zebras lowering their head and waving it from side-to-side (its strongest form of threat), elephants swaying and nodding their head, and chimpanzees swaying from side-to-side [36]. Even we, ourselves, use the simple up-and-down and side-to-side motions as responsive ‘Yes’ and ‘No’ head gestures. These motions can also be easily found in other non-communicative behaviors.

These simple motions are also not necessarily restricted to movement of a specific body part or limb. For example, various birds perform the up-and-down display movements using the head, wing, body, foot, and tail (See Col. 1, Table 3). The side-to-side motions can also be readily observed in the head, body, and tail of birds (See Col. 2, Table 3). This diversity of limb and body

performance for the motions suggests a higher-level, more “cognitive”, control structure for the movements [68].

| <b>Table 3. Diversity of avian anatomy used for performance of motions.</b> |                        |                          |
|---|------------------------|--------------------------|
| <b>Body Part</b>  | <b>Up-and-down</b>     | <b>Side-to-side</b>      |
| Head  | Mallard duck [94]      | Pileated woodpecker [32] |
| Wing  | Ostrich [10]           | –                        |
| Body  | Bluejay [94]           | River kingfisher [44]    |
| Foot  | Black guillemot [33]   | –                        |
| Tail  | Belted kingfisher [95] | Mockingbird [94]         |

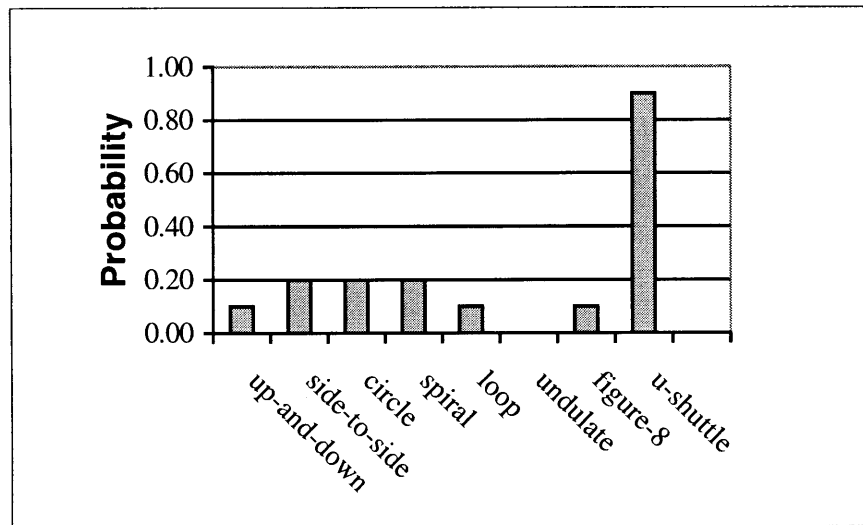
When we examine the histogram distribution of the oscillatory display motions as seen across 14 types of birds (ducks, falcons, gulls, hawks, herons, hummingbirds, kingfishers, larks, owls, pigeons, plovers, terns, wheatears, and woodpeckers as described in a dozen field guides [16, 17, 32, 33, 90-92, 94-97], See Appendix A), we see a biased trend favoring the *simplest* movements in the hierarchical ordering (See Figure 11). Although most bird species can perform many of the movements to some degree, the relative frequency of these motions differs considerably *across* species. For example, the up-and-down, side-to-side, and circle motions are quite ubiquitous across species, and not surprisingly, the more complex motions such as figure-8, U-shuttle, and loop are more infrequent. To test for a precise correlation between the ubiquity across bird species and our hierarchical organization (under the assertion that “simpler movements should be more universal”) would require a thorough examination of all bird species and a comprehensive cataloguing of their behaviors – a daunting task. Our sampling of bird species, though, indicates that there indeed is a general trend: simpler oscillatory movements (1-D, 2-D) are more widespread across species than are the more complex patterns (2-D + translation, 2-D with frequency-doubling).



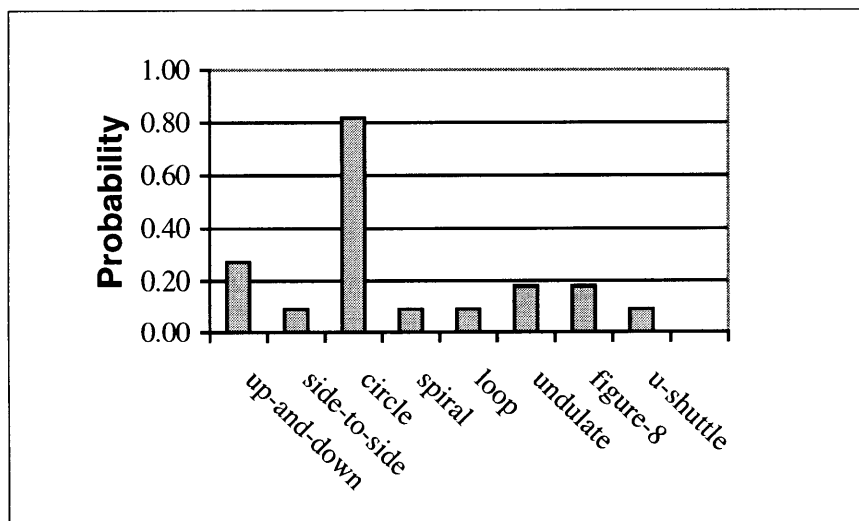
**Figure 11.** Motion generality for birds. Overall motion preference across 14 types of bird (e.g. ducks, hawks, larks).

But when we examine these motions for various *individual* bird species, we see that different types of birds have a tendency to perform certain characteristic motions. This is shown in more detail in the histograms presented in Figure 12 - Figure 14, where the peaks correspond to the most common oscillatory movements performed by specific species of bird (hummingbirds, falcons, and woodpeckers). Here we see that most hummingbirds perform the *U-shuttle* (See Figure 12), falcons tend to fly in *circular* formations (See Figure 13), and woodpeckers particularly move *up-and-down* and *side-to-side* (See Figure 14). The data was compiled by examining the oscillatory displays of several species for the three bird types (10 hummingbirds, 21 woodpeckers, and 11 falcons, See Appendix B) as observed in the field guides [17, 32, 33, 91, 94-97]. For each type of bird, the species count of each motion was normalized by the number of species examined to give an estimated probability that the motion will be exhibited by that type of bird. Individual preferences for the oscillatory motions show that not all motions are equally preferred by different types of bird. But as birds do not equally use the motions for their displays, the *ubiquity* of certain patterns (1-D, 2-D) *across* species as given above does show evidence for our notion of simplicity. Thus we believe due to the prevalence of the simpler motions across species (different animals or across birds) and their limb diversity for performance, that the up-and-down and side-to-side (and perhaps circle) motions are a reasonable prediction for the simplest motions and therefore support our notion of “*simplicity as ubiquity*”.

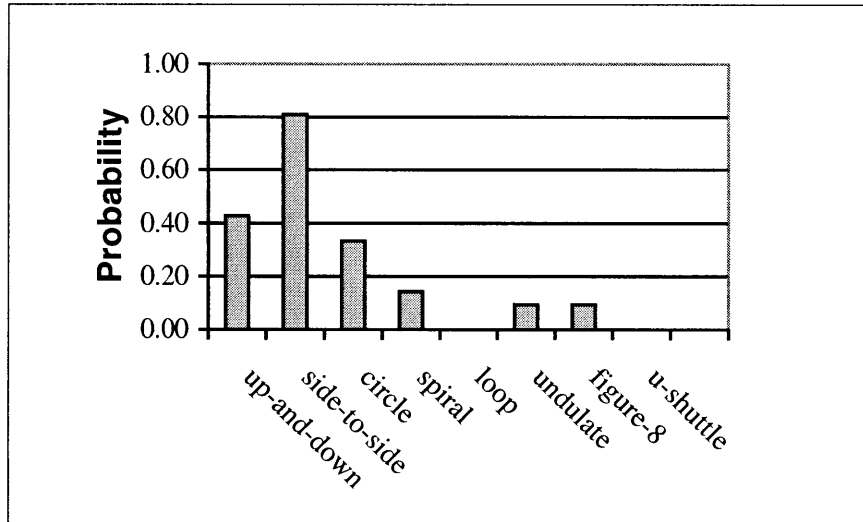




**Figure 12.** Hummingbird preferences for U-shuttle motion from examining 10 different species.



**Figure 13.** Falcon preferences for circle motion from examining 11 different species.

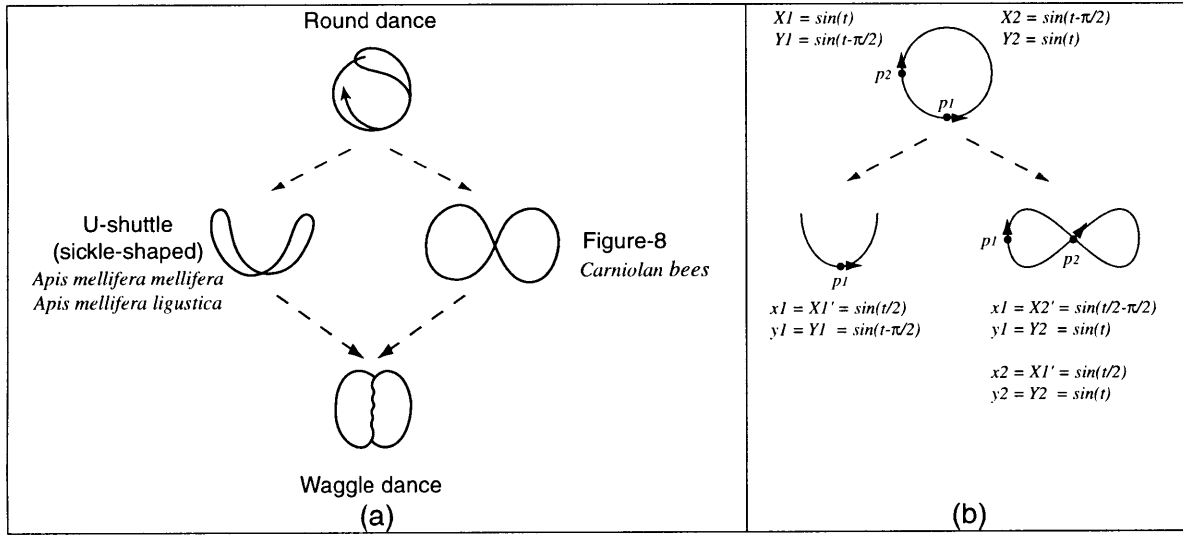


**Figure 14.** Woodpecker preferences for up-and-down and side-to-side from examining 21 different species.

### 3.4.2 Dance language of bees

We also find that the model-ordering relationship for circle, figure-8, and U-shuttle motion (See Figure 8) is similar to the transformational relationship in the communicative dance language of honey bees. It is well known that bees perform particular stylistic dances in the hive to signal the presence and location of a food source to others [108]. In the *round* dance, the bee swiftly runs in circles, suddenly reversing direction after the circular march (often after one or two complete circles). This dance signals to the other bees that there is a nearby feeding source up to 10 meters away. When food is beyond 100 meters, the bee performs a *waggle* dance. In this dance, the bee runs straight ahead for a short distance, returns in a semicircle, again runs through the straight stretch, returns in a semicircle in the opposite direction, then repeats the entire pattern. During the straight part of the run, the body vigorously wags back-and-forth (13-15 Hz). This dance trajectory is similar to a figure-8 pattern and conveys the distance to the food source (by the dance tempo) and the compass direction to the site (from the direction of the straight run). Mechanical models performing the waggle dance have even shown success at eliciting the appropriate behaviors of honey bees [72]. From a purely informational perspective, these communicative dances correspond to simple (round dance) and complex (waggle dance) signaling mechanisms. Interestingly, when the food source is moved from 25 to 100 meters, a transition between the round and waggle dance occurs. As the distance to the food source is increased past 25 meters, Carniolan bees transform from the circle pattern (round dance) to the path of a pure figure-8, then into the classic waggle dance. With other races, for example *Apis*

*mellifera mellifera* or *A. m. ligustica*, the transition involves moving in a sickle-shaped (U-shuttle) pattern rather than the figure-8 before transforming into the waggle dance. The dance movements and the two possible transitions are depicted in Figure 15. With there being fewer parameters needed to generate the circle dance as compared with the more complex waggle dance, we interpret the information content and transitional pathways for the dances as a reflection of a hierarchical complexity ordering for the circle, figure-8, and U-shuttle motions. Interestingly, this biological ordering for the three motions is identical to the earlier model-based ordering for these patterns with increasing specialization in the parameters corresponding to a frequency-doubling constraint (See Figure 15). Thus we see a strong correlation between a biological organization for the motions and our hierarchical specialization based on model parameter complexity.



**Figure 15.** Transformational ordering of dances for honey bees as the distance to a food source is moved increasingly farther. (a) Carniolan bees (right side) dance in a figure-8 pattern between the round dance and the waggle dance. Other species (left side) instead perform a U-shuttle sickle-shaped dance in place of the figure-8 during transformation. (b) Model ordering based on parameter specialization with frequency-doubling. This ordering is the same as shown in (a) for the three motions.

### ***3.5 Section summary***

In this section we demonstrated in detail the characterization, organization, and biological relation for a set of categorical oscillatory motions. A simple sinusoidal model with very limited and specific parameter values was presented that characterized the oscillatory motions with both structural and stylistic components. The motions were then ordered from simple to complex in a hierarchical organization by the number of model parameters and constraints. With this ordering, the primary sub-categories in the motion set were derived. This organization was shown to have a general correlation with observations of bird and honey bee movements. The significance of this categorical organization for the oscillatory motions in terms of computation is described in the following section, where the sub-categorical ordering is used to guide the design and layout of computation for recognizing these patterns.

## 4 Computational Implementation

The goal of this research is to exploit the meaningful natural constraints in *categorical* motion patterns for the development of a categorical *recognition* model. Our computational approach to the recognition of the oscillatory motions is to extract the relevant sinusoidal parameters from the input data using computations that reflect the ordered sub-categories derived for the motions (in the previous section). With the sinusoidal model, we have shown a simple to complex ordering of four sub-categories (1-D, 2-D, translatory, and frequency-doubling motion). We interpret these groupings to reflect multiple processing levels within a computational recognition framework. Therefore, a basic-level substrate is first developed to recognize the simplest patterns (1-D motions), and is then evolved to full categorical recognition by adding successive computations for the additional specializations needed to identify the more complex motions (2-D, translation, frequency-doubling). We first segment a single pattern cycle from the input position trajectories using sign-changes (zero-crossings) in the velocity trajectories to find the starting and ending locations of the pattern. Second we extract the required oscillatory motion parameters (of amplitude, frequency, phase, translation) from the x and y trajectories using Fourier analysis. Depending on which motion or sub-category is to be recognized, certain sub-categorical computations are enabled to estimate the motion parameters and match them with the specified parameterized models.

In our implementation, we rely on the 2-D projected position information for recognition. Not to say that 3-D motion is not of consequence, but we accept the 2-D projection to be sufficient for recognition of these motions since all but one of the motions are at most 2-D and the structural nature of the motions persists through non-extreme projection or viewpoint (i.e. projections of the oscillatory motions from the world to the image retain their oscillatory structure). For example, a circle projects into either a circle, ellipse, or the degenerate case of side-to-side motion. The projection of the more complex figure-8 pattern remains a figure eight, unless positioned in the degenerate case where it projects into a flat oscillation.

Here we present the methods used for implementing the categorical recognition based on the previous categorical organization for the motions. We discuss the details for the data representation, lowpass filtering, pattern segmentation, parameter estimation, computational sub-categories, and periodicity calculation. Before describing the implementation, we give a brief

discussion of related work in computer vision pertaining to the recognition of periodic and oscillatory motion.

#### **4.1 Related work**

There is much related work by computer vision researchers on recognizing periodic and oscillatory motion. Heavy reliance in these techniques is placed on Fourier analysis of trajectory information to detect and recognize periodic motion, while others examine specialized representations outside of the frequency domain.

For example, Polana and Nelson [79] employ a template-based method registering periodicity values within an overlaying grid on the video-sequence. They recognize people walking in outdoor scenes by gathering a feature vector, over the entire body, of low-level motion characteristics (optical-flow magnitudes) and periodicity measurements. They use repetitive motion as a strong cue for recognizing cyclic walking motions. After gathering multiple training samples, recognition is performed using a nearest centroid algorithm. Similarly, Liu and Picard [62] use periodicity templates as a representation to localize and detect repetitive motion along temporal lines for each pixel region. The method first stabilizes the object in motion, aligning the object to the same position in each frame. Then 1-D signals along the temporal dimension for pixels in the X-Y-T image volume are decomposed into deterministic (periodic) and indeterministic (random) components using the Wold decomposition. The ratio of harmonic energy to the total energy is used to detect and measure the periodicity of the signal. Cutler and Davis [27] check a self-similarity measure between motion frames for detecting periodic motion. First, they segment the motion and track the object throughout the sequence. Next, the object is aligned along the temporal axis, and a self-similarity metric is calculated between the images of the tracked object. This temporal measurement yields a periodic signature if the motion itself is periodic. Their technique is different than the above methods in that they analyze the periodic similarities of larger object areas, not at small or individual pixel regions. Tsai et al. [103] detect cycles in a walking pattern using autocorrelation and Fourier transform techniques of a spatio-temporal curvature function from point trajectories. For each trajectory, a large impulse in the Fourier magnitude plot indicates the cycle frequency. Recognition is achieved by matching one cycle of the input trajectory to a model cycle using an array of zero-crossing potentials. Other research targeting periodic walking patterns include work by Little and Boyd [61] which

examines the phase relationships of periodic elements derived from optical flow for recognition and person identification. They recognize people walking by analyzing the motion associated with two ellipsoids fit to the body. One ellipsoid is determined using the motion region silhouette of the person, and the other ellipsoid is calculated using motion magnitudes as weighting factors. The relative phase of various measures (e.g. centroid movement, weighted centroid movement, torque) over time for the ellipses was used to characterize the gait of several people.

Non-Fourier approaches to cyclic motion analysis include Niyogi and Adelson's approach [77] of detecting braided walking patterns in spatio-temporal slices. Their method is based on the observation that walkers generate special signatures in *space-time*. The approach models the gait pattern of the person as a set of braided contours in the X-T slice of the total X-Y-T image volume. They analyze these patterns using snakes and use them to estimate the model parameters of a simple moving stick-figure. For recognition, the walking signals are time-warped and matched using a simple k-nearest neighbors measurement. Seitz and Dyer's method [89] present an analysis of periodic motion using a period trace and view-invariant image matching. The period trace maps of all the cycles in a repeating pattern and can be used to identify motion trends and irregularities. The approach also uses an affine invariant match function to compare images and determine whether an image sequence is the projection of a true 3-D cyclic motion. As with many other image-based methods, the approach is applicable to non-rigid cyclic motions. Cohen et al.'s technique [23] uses linear dynamic models to recognize simple oscillations from flashlight point data. Their approach is motivated with a simple set of gestures that include oscillating planer motions (circles and back-and-forth lines). A differential equation is used to represent the motion dynamics of the gestures and with different parameter settings acts a predictor between different oscillations. Recognition identifies the model with the smallest prediction residual.

We continue the development of approaches for the analysis and recognition of repetitive motion, but instead offer a categorical method based on strong structural regularities found within a set of naturally occurring, animate oscillatory motions. Our goal is not to identify periodic motion (as defined earlier), but is to recognize particular motion patterns that are of a specific sinusoidal and oscillatory nature. Before presenting the implementation, we discuss an experiment used to determine the form of the data representation.

## **4.2 Data representation**

Before constructing the motion recognition system, a representation for the data must first be chosen. We require the data to be time-based, rather than purely spatial, due to the temporal nature of the signal and as a result of experimental findings for the speed and frequency tunings involved in perception for specialized recognition. However, it is not obvious whether the patterns should be characterized using position, velocity, acceleration, etc. The answer gives us some insight into the form of the generative model for the motions. Given the choice between data representations for the oscillatory motions, we select the one that best retains a sinusoidal shape over varying performance conditions. We examine specifically the trajectories of human arm motion. For our oscillatory motion study we looked at human arm performance because it is convenient for experimentation (difficult to collect extensive videos of animal oscillations) and it will be used as the primary mode of interaction for later experiments and demonstrations. Our analysis of representational form based on purely human performance will therefore be biased toward our own motor capabilities.

Previous experiments show that the speed profiles of reaching arm movements depict a bell-shaped form. A number of models based on “minimum-jerk” principles have been proposed for such arm/hand movements (e.g. [39, 73, 102]), where this model predicts straight paths and bell-shaped speed profiles matching the experimental results. For a given path, the general model assumes that the correct speed profile will be the one that minimizes the third derivative of position (“jerk”) using average data. For example Viviani and Flash [107] extracted velocities from hand trajectories and used optimization methods to produce the minimum-jerk path and speed profile. Atkeson and Hollerbach [5] showed that unrestrained human arm trajectories between point targets have an invariant velocity profile when normalized for both speed and distance. Movements in this case were executed and tracked between different points under varying conditions of speed and hand-held load. The interpretation of small high-order derivatives is that of smoothness in the path, resulting in more “predictable” movements [50]. These experiments all use simple and generic point-to-point reaching, pointing, and turning movements (some closed parameterized curves) of the arm. We will show that arm performance of the cognitively-based oscillatory patterns (rather than positional movement tasks) show considerable derivative variance whereas the position information remains more stable. Thus for



these patterns, the variance in the overall position, rather than derivative, appears to be minimized (i.e. minimizing jerk may not be appropriate).

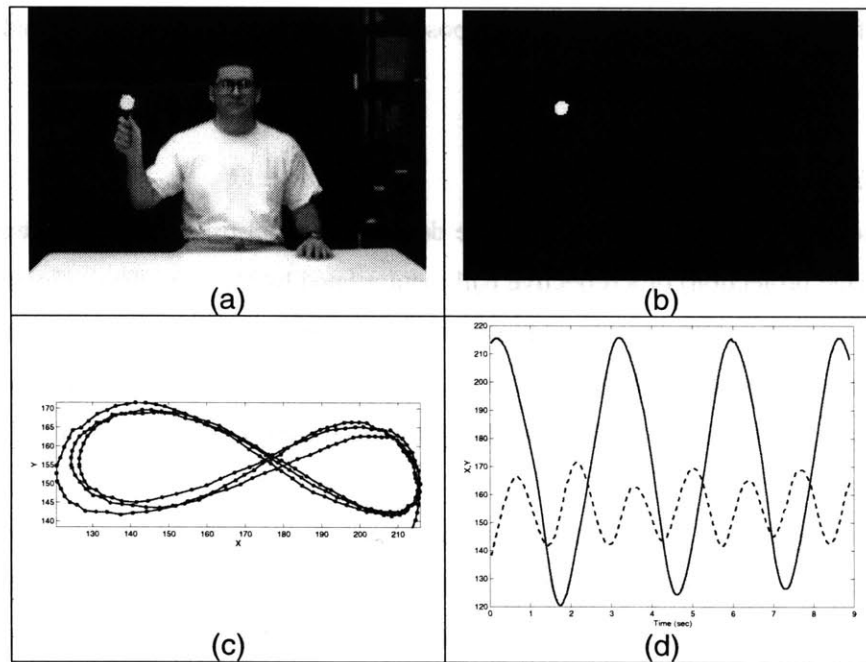
In a similar fashion to these experiments, we examined human-generated patterns of oscillatory motion with alterations in user, speed, amplitude, and performing limb. The variations allowed us to measure the stability and invariance for both position and velocity. The position and velocity trajectories are time-based and both contain the parameter information necessary for the sinusoidal model characterization (the sinusoidal parameters persist through the velocity derivative calculation). In [5], an LED system was employed to accurately track the arm positions. To conduct our experiments, we developed a similar real-time tracking mechanism that automatically and robustly extracts the 2-D position trajectories of a reflective ball manipulated by the user.

#### *4.2.1 Point-light display*

To aid in the study of human arm motion, we developed a real-time system that extracts the 2-D position (image projection) of a reflective ball manipulated by the user. The reflective ball system was created to ease in scene segmentation and feature extraction, where the task at hand is to concentrate on the “shape or path of motion”, rather than on feature identification followed by tracking. These trajectories of the ball are then be used to analyze the motion of the arm in detail.

Our method employs an infrared light source (840 nm), an infrared-sensitive black-and-white video camera (Sony SSCM370) equipped with an infrared bandpass filter (optically clear from 700 to 1100 nm), and a wand with a light-reflecting ball attached to the top of the shaft. The infrared light source and filtered video camera are directed at the user who is approximately 3 meters away from the setup (See Figure 16.a). The brightness gain on the video camera is lowered to register the infrared light reflected from only this highly-reflective ball. Therefore, any other visibly bright objects in the image that do not significantly reflect or emit infrared light (e.g. the shirt and desk in Figure 16.a) are optically filtered out of the signal. A computer system then digitizes the video input, thresholds the images (See Figure 16.b), extracts the centroid of the 2-D ball region in each image, and records the positions. We accept the 2-D projection of the ball motion to be sufficient for comparing the position and velocity representations in this study. If required, the depth changes could also be roughly approximated from the point-light displays by calculating the changes of the projected ball radius in the image.

With this system, a person can freely maneuver the wand while the imaging system extracts the 2-D image location of the reflective ball. For these experiments, the video is digitized and processed off-line (though the entire system can run in real-time at 30 Hz on a Silicon Graphics R10000 O2 workstation). The design and output of this system resemble the optical motion capture systems used for motion analysis and animation ([4, 11, 28, 55, 71, 86]). Given the puppetry-style of the setup, it can be expected that certain dynamical properties of animal motions are not adhered to fully by the arm motion (e.g. arm control may not correctly mimic the dynamics of full body motion of other creatures). Example data trajectories generated with the system for a figure-8 pattern by the user are shown in Figure 16.c,d.



**Figure 16.** Infrared ball tracking system. (a) Person holding a wand attached to a light-reflective ball. (b) Ball region extracted by vision system. (c) Example user-generated spatial pattern. (d) Data trajectories. (Solid line: x-trajectory, dashed line: y-trajectory).

#### 4.2.2 Experiments

The first of four experiments comparing the position and velocity of human arm oscillations was comprised of having three individuals each perform multiple repetitions of a sideways figure-8 pattern (requiring sinusoidal movement in both  $x$  and  $y$ ). Figure 17 shows the layered normalized plots (in both time and amplitude) of successive half-cycles for the mean-shifted position (left) and velocity (right) trajectories. Each generated data trace is initially lowpass filtered to remove any noise that could adversely affect the derivative calculation but not to the

point of modifying the true waveform shape. As shown in Figure 17, the  $\dot{x}(t)$  velocity profile shows considerable variance while the  $x(t)$  position trajectory remains stable across users. The  $y(t)$  and  $\dot{y}(t)$  plots are more consistent with one another, and remain so throughout the following experiments. The second test involved the same three users performing the motion, but with each performing the sideways figure-8 motion at two different speeds (slow and fast). Still,  $x(t)$  remains more stable than  $\dot{x}(t)$ , discounting the one anomalous half cycle (See Figure 18). The third test had the users perform the motion pattern at two different amplitude heights (vertically enlarging the loops). Again, as presented in Figure 19, a consistent  $x(t)$  and variable  $\dot{x}(t)$  are found. The final test examined the motion of one person performing the motions using the upper body, arm, and forearm. For the upper body motion, the user donned a hat with a reflective ball attached, leaned forward, and swung the torso in the figure-8 pattern. The arm motion was generated by holding and moving the wand while keeping the elbow joint locked. Lastly, the forearm motion was performed while holding the upper arm tightly against the body. As shown in Figure 20, the position data is once again more consistent than the velocity data. For all the tests, the position data basically retained the same characteristic shape.

Naturally, taking derivatives of the position signal to generate the velocity profile introduces noise into the signal. This variance increase from the derivative operation corresponds to a non-linear multiplier of the frequency magnitude (increasing to the Nyquist frequency):

$$\dot{x}(t) = \frac{x(t) - x(t - \Delta t)}{\Delta t} \rightarrow \sqrt{\frac{2(1 - \cos(2\pi f \Delta t))}{\Delta t^2}} |X(f)| \quad (2)$$

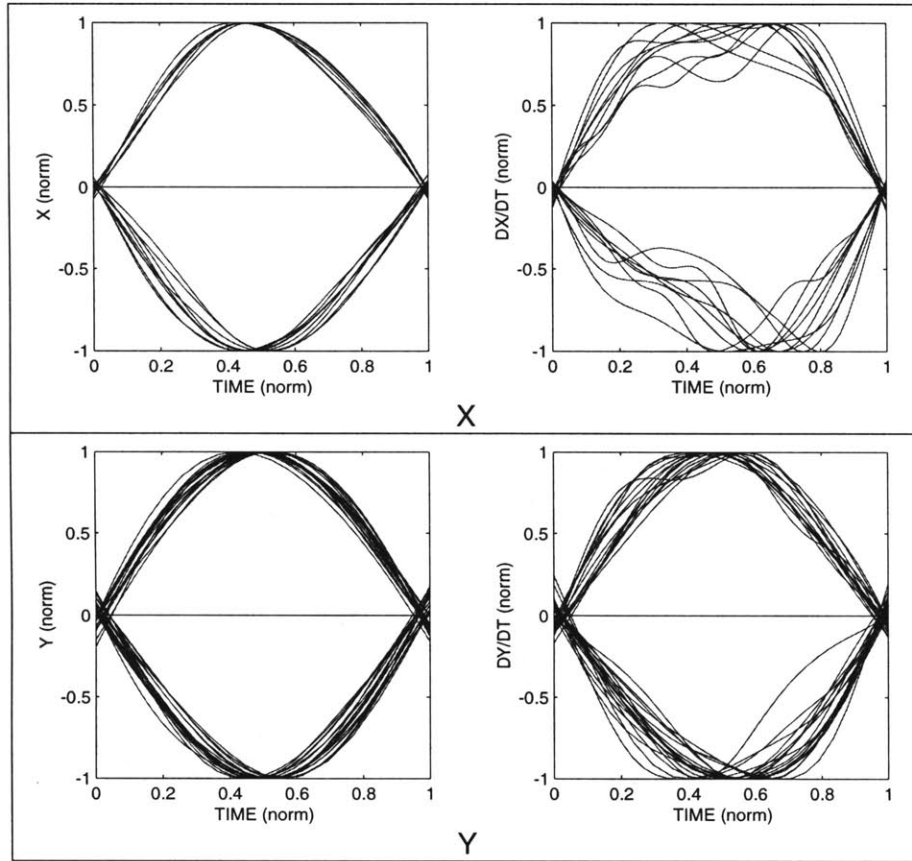
With real-time processing,  $\Delta t$  is typically 0.0333 seconds (30 Hz). As the derivative approximation becomes the true derivative (i.e.  $\Delta t \rightarrow 0$ ), the multiplier, using L'Hôpital's rule, becomes a linear ramp as a function of frequency:

$$\lim_{\Delta t \rightarrow 0} \sqrt{\frac{2(1 - \cos(2\pi f \Delta t))}{\Delta t^2}} |X(f)| = 2\pi f |X(f)| \quad (3)$$

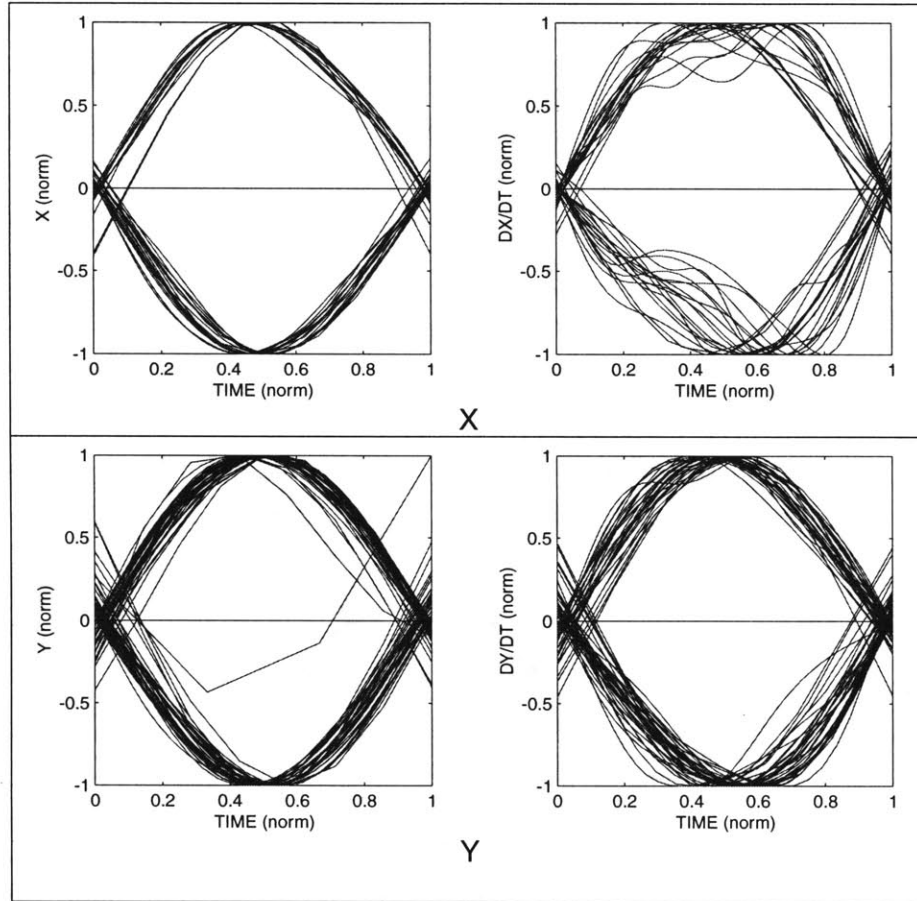
Thus the derivative of the original signal has increased noise regardless of the sample rate. Because naturally generated sinusoid motions have more variation and typically have harmonic components, the resultant derivative contains significantly amplified noise. Also, the increase in

horizontal variance in the velocity (See  $\dot{x}(t)$  vs.  $\dot{y}(t)$  in Figure 17) may be partially accounted for by the higher degrees of freedom in the arm for side-to-side motion (elbow and shoulder) than for up-and-down motion (hinged elbow). This variance is reduced in Figure 20 where the motion of the elbow joint was minimal.

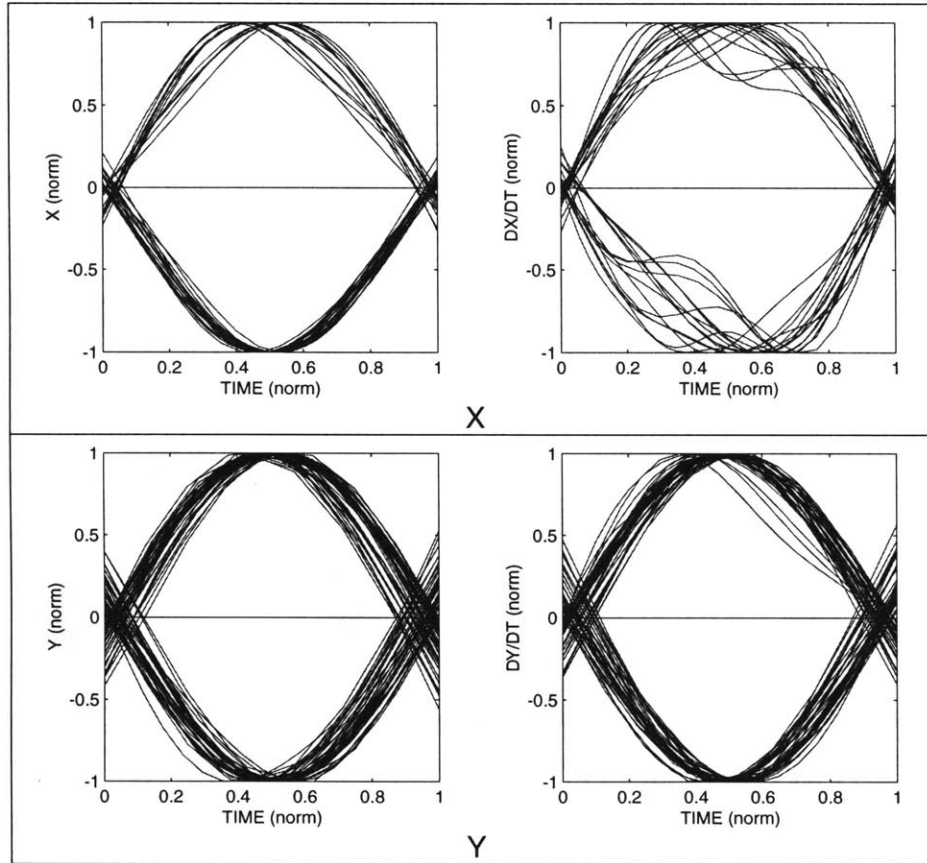
These experiments suggest that we choose the more stable position trajectories, and not the velocity, as the data representation for the analysis and recognition of the motion patterns (for arm motion). Though we could theoretically cancel the noise from the velocity trajectories, it is more complicated, requiring additional steps, and is likely not to be absolute. This promotion of position over velocity is interestingly contradictory to the previous derivative-based models.



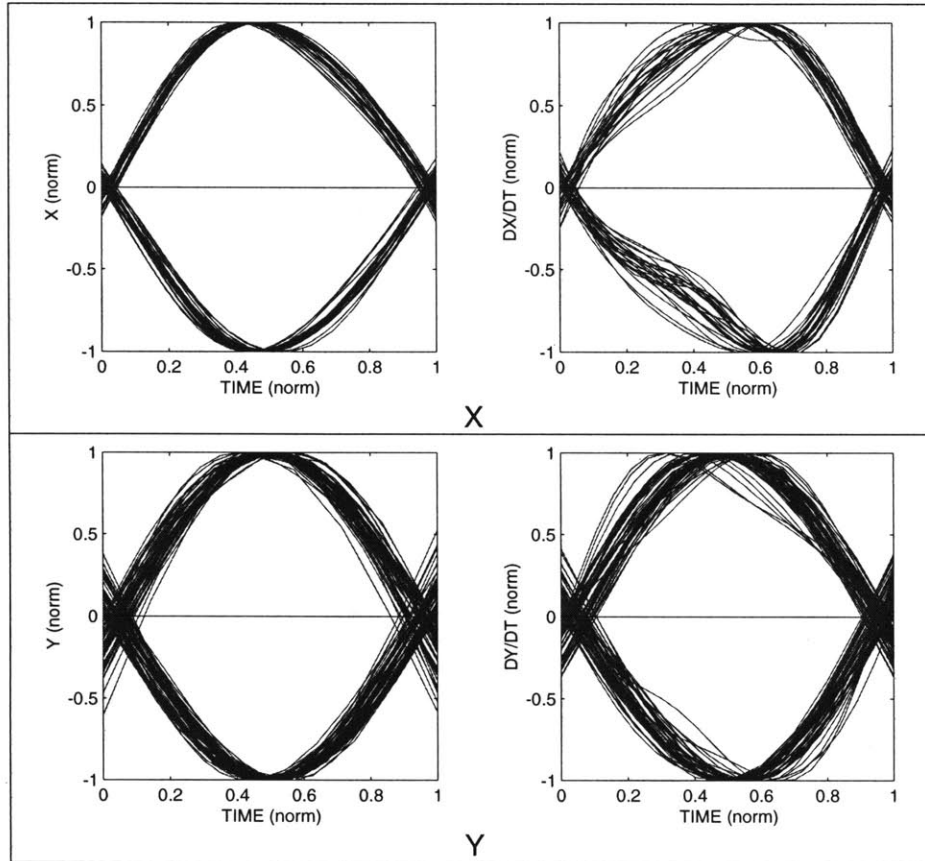
**Figure 17.** Experiment 1. Composite half-cycles for three users performing a sideways figure-8 motion ( $\infty$ ). Top row shows normalized  $x(t)$  (left) and  $\dot{x}(t)$  (right) motion trajectories. Bottom row shows normalized  $y(t)$  (left) and  $\dot{y}(t)$  (right) motion trajectories.



**Figure 18.** Experiment 2. Composite half-cycles for normalized  $x(t), y(t)$  (left) and  $\dot{x}(t), \dot{y}(t)$  (right) figure-8 motion trajectories for three people at two different speeds.



**Figure 19.** Experiment 3. Composite half-cycles for normalized  $x(t), y(t)$  (left) and  $\dot{x}(t), \dot{y}(t)$  (right) figure-8 motion trajectories for three people at two different height magnitudes.



**Figure 20.** Experiment 4. Composite half-cycles for normalized  $x(t), y(t)$  (left) and  $\dot{x}(t), \dot{y}(t)$  (right) figure-8 motion trajectories. Data shows results of the head, arm, and forearm motion of a single person.



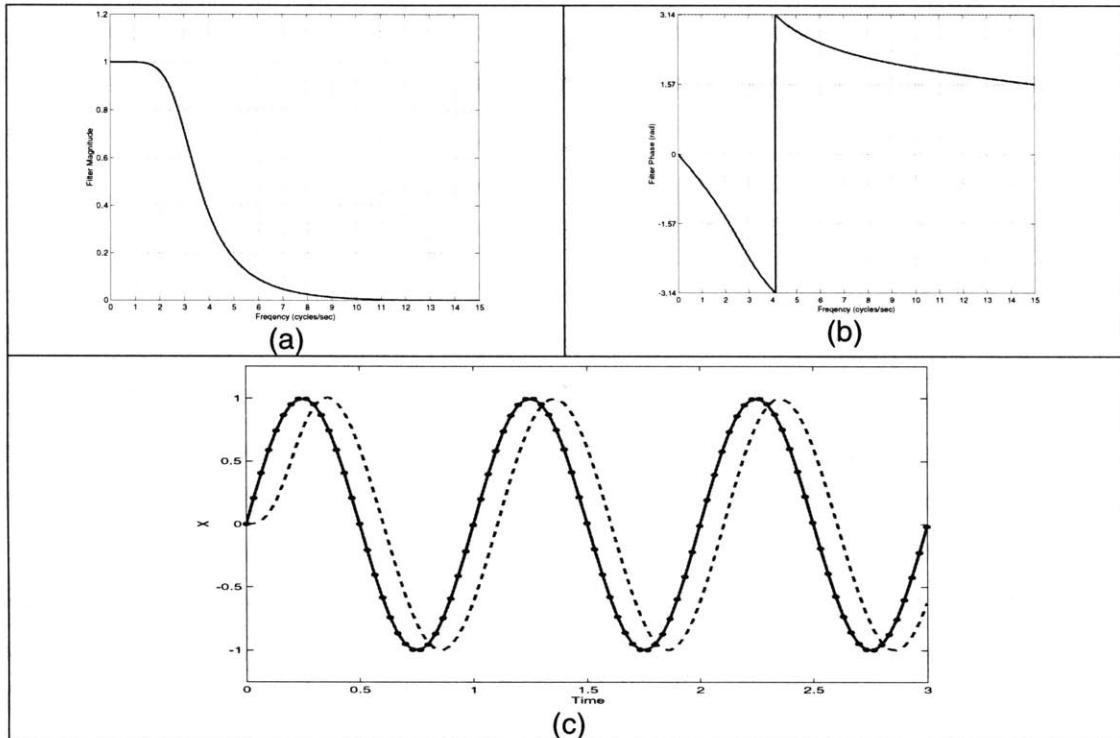
## 4.3 Methods

### 4.3.1 Lowpass filtering

Initially, the raw input signal of a tracked object or feature is filtered to remove noise resulting from the extraction of the point location  $(x(t), y(t))$ . We chose an IIR Butterworth lowpass filter to smooth the data due to the simplicity of computation required using a small number of filter coefficients. The filter is an implementation of the standard difference equation:

$$\begin{aligned} a(1)*\hat{x}(n) = & b(1)*x(n)+b(2)x(n-1)+b(3)*x(n-2)+b(4)*x(n-3) \\ & -a(2)*\hat{x}(n-1)-a(3)*\hat{x}(n-2)-a(4)*\hat{x}(n-3) \end{aligned} \quad (4)$$

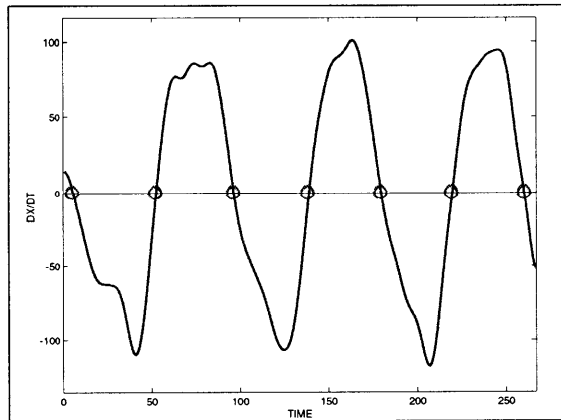
with  $a(1)=1.0000$ ,  $a(2)=-1.7600$ ,  $a(3)=1.1829$ ,  $a(4)=-0.2781$ ,  $b(1)=0.0181$ , and  $b(2)=0.0543$ ,  $b(3)=0.0543$ , and  $b(4)=0.0181$  for a 3<sup>rd</sup> order lowpass Butterworth filter with a cutoff frequency of 3 Hz at a 15 Hz Nyquist rate. We chose this frequency passband because the motions used in the later experiments fell mostly within this range. Figure 21.a,b shows the magnitude and phase response plots for this particular filter, and Figure 21.c shows the results of filtering a 1 Hz sinusoid. One disadvantage of an IIR filter is that it gives a non-linear phase response that delays different frequencies by different amounts (unlike an FIR filter). In the case of the Butterworth filter designed for this research, the phase in the passband is fairly linear (3.6 degrees SD from linear scaling function of frequency between 0 and 3 Hz). An upper bound on the phase difference error in our design is calculated for two filtered signals (with one signal having twice the frequency as the other, as in figure-8 motion) to be 18 degrees (e.g.  $\max(2\Delta_{slow}-\Delta_{fast})$ , where  $\Delta$  is the non-linear phase error from the linear fit). We will use this error bound as a single tolerance measure during recognition for all motions. Though we could perform forward-backward filtering with this same filter to achieve linear phase, we opted for the simplest forward-only approach.



**Figure 21.** Butterworth digital lowpass filter (cutoff at 3 Hz, 15 Hz Nyquist rate, coefficient order of 3). (a) Magnitude response. (b) Non-linear phase response. (c) Result of applying filter to sinusoid input (frequency=1 Hz, phase = 0 rad). Solid line is original data and dashed line is filtered data. The filtered data has the same amplitude and frequency as the original data, but is phase shifted by  $-37$  degrees or 0.1029 sec (approximately 3 samples).

#### 4.3.2 Pattern segmentation

We designed the system to require only a single pattern cycle in the x and y position trajectories to recognize the pattern. Many other computer vision methods require extended cycles for periodicity recognition. We feel that the single pattern cycle is the salient motion pattern rather than an extended sequence of the patterns. One way to find a single pattern cycle is to look for a certain number of peaks and valleys in the sinusoidal position trajectories. Peak/valley detection may not be reliable because *multiple local* measures are needed to find an extrema location. A better approach is to find features that are uniquely defined given only one or two robust measurements. Sign-changes, or zero-crossings, in the corresponding velocity trajectory (See Figure 22) satisfy this objective and can be reliably used to similarly extract the pattern cycle. These zero-crossings correspond to the locations of the peaks and valleys in the position trajectory. A single pattern cycle is then found by counting a certain number of sign changes in velocity within a designated trajectory. Though the velocity was determined earlier to be more susceptible to magnitude noise than position, its zero-crossings can be robustly identified to segment the pattern. After the pattern cycle is extracted, the motion pattern is shifted by the pattern center to remove the viewing condition. Recognition therefore occurs at these velocity zero-crossings in the pattern rather than continually throughout the motion or only once at the end of a complete cycle.



**Figure 22.** Velocity zero-crossings of mean-shifted x-position trajectory for a figure-8 pattern.

#### 4.3.3 *Parameter estimation*

We use Fourier analysis to estimate the desired sinusoidal model parameters (amplitude, frequency, and phase) from the extracted pattern cycle. The 1-D Fast Fourier Transform (FFT) is used to quickly and reliably estimate the required sinusoidal parameters individually for the x and y trajectories. The Numerical Recipes FFT implementation [80] is used directly with no windowing of the data (e.g. Hanning). The parameters are estimated for each trajectory in isolation from the other trajectory and later combined and compared for the structural analysis. Later in the experimental stage, we examine this method of parameter estimation with varying types of signal noise.

#### 4.3.4 *Computational sub-category levels*

We interpret the constrained parameter relations (Table 2) and sub-categorical complexity ordering of the motion patterns (Figure 10) to reflect a series of localized and layered processing components within a fixed category for computational recognition. An initial substrate is first developed to recognize the simplest 1-D motion patterns (up-and-down, side-to-side), and is then evolved to full categorical recognition by successively adding computation in layers to accommodate the increasing complexity divisions of 2-D (circle), translation (spiral, undulate, loop), and frequency-doubling (figure-8, U-shuttle). Therefore, the amount and type of processing for recognizing a motion depends on the complexity of the pattern to recognize. This layering of computation is similar to the robotics Subsumption architecture championed by Brooks [18], where complex behaviors are developed by successively layering independent computations that affect the input and output of new and existing computational layers.

Each motion model in the category consists simply of its sub-category and specific values for particular sinusoidal model parameters or relationships. The job of recognition is therefore to extract and match the parameters values from the input data to the model (using the appropriate sub-categorical computations). For matching the parameter values and constraints, the method can employ two separate tests. The first test examines the structural relations for the parameters, which may include (depending on the computational level) amplitude ratios, frequency ratios, phase differences, and translation presence. This test verifies that the categorical *structure* is adhered to in the input data. After a correct structural match has been found, a selective *stylistic* examination (if required) determines if the actual stylistic parameter values match the specific values in the model. This test can check to see if specific amplitude values (e.g. the height) or

frequency values (i.e. the speed) are the necessary quantities required by a specifically-tuned model.

To calculate the amplitude, frequency, and phase parameters (when each is applicable) for the pattern cycle, Fourier analysis (FFT) is applied to the input data trajectories. An amplitude threshold (3 pixels) is then compared with the data to remove small motion trajectories, and a *relative-amplitude* scale threshold (10%) is applied to remove a small motion trajectory overshadowed by a reference sinusoidal trajectory (the most “stable” trajectory determined from the motion to recognize).

The trajectory to be examined for the velocity zero-crossings (for pattern segmentation) and the extent of model parameters to be estimated and matched are determined by the sub-category of the target motion. We now describe the sub-categorical organization and procedure for those computations.

#### 1-D motion: up-and-down, side-to-side

The basic substrate for recognition is designed to recognize the primitive 1-D oscillatory motions. The position data for both trajectories are extracted between the first and third velocity zero-crossings of the sinusoidal trajectory (one pattern cycle). Which trajectory is examined for the sinusoid is determined by the model being considered for recognition. For up-and-down, the sinusoidal trajectory is the vertical y trajectory; for side-to-side the sinusoidal trajectory is the horizontal x trajectory. The structural test for these motions confirms that only the sinusoidal trajectory has motion (a non-zero amplitude). The stylistic test can match the actual frequency and amplitude measurements to model specifics.

#### 2-D motion: circle

In the second sub-category level for 2-D motion, the pattern is extracted between three velocity zero-crossings, all in the x or the y trajectory. The analysis follows the same pathway as with 1-D motion, but additionally verifies motion in *both* trajectories (both amplitudes must be non-zero), that their frequency ratio is unity ( $f_x = f_y$ ), and that their phase difference  $\phi_x - \phi_y$  is  $\pm\pi/2$  ( $\pm 18$  degrees calculated as the maximum error from the non-linear lowpass filter response). The stylistic test verifies the singular frequency value as for 1-D motion, but additionally checks the amplitudes for *both* trajectories.

#### Translatory motion: spiral, loop, undulate

Movements that have a translation component are incorporated in this next sub-category. Three velocity zero-crossings in the non-translating trajectory, known a priori from the model to be recognized (y trajectory for loop and undulate, x trajectory for spiral), are used to extract the pattern cycle. To remove the translation, the linear component between the start and stop points of the translated trajectory is subtracted. The structural test for the amplitude, frequency, and phase are the same as for 2-D motion, though not all parameters and tests are used for the single undulate motion (the frequency ratio and phase difference checks are suppressed for the undulating motion due to a translation-only trajectory). An additional check is made for a translation value above (5%) a relative minimum (relative to the translating trajectory amplitude for spiraling and looping, or relative to the non-translating trajectory amplitude for undulating). Otherwise no translation is identified. The stylistic test as before checks the amplitude and frequency values and may additionally look for a specific translation quantity according to the model specifics.

#### Frequency-doubling motion: figure-8, U-shuttle

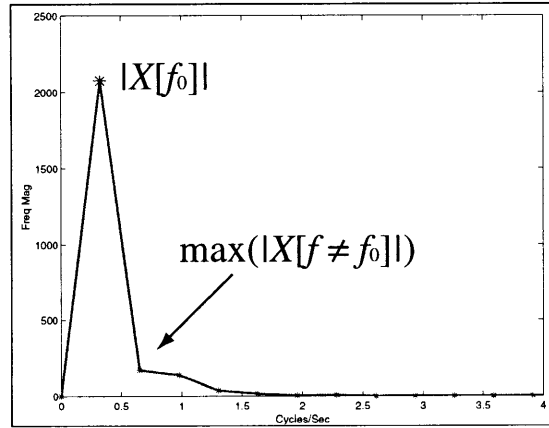
In this last sub-category, five velocity zero-crossings are found in the trajectory with the faster frequency to extract a single pattern cycle (earlier experiments show the faster trajectory to be more stable). In the sideways figure-8 and U-shuttle, the faster trajectory is the up-and-down y trajectory. In the structural test, both trajectories must have motion with a change in the frequency ratio to one half ( $f_x = \frac{1}{2}f_y$ ), and a relative phase difference ( $2\phi_x - \phi_y$ ) of 0 or  $\pi$  ( $\pm 18$  degrees) for figure-8 motion or  $\pi/2$  ( $\pm 18$  degrees) for U-shuttle motion. The stylistic tests remain the same as in 2-D motion.

#### *4.3.5 Periodicity calculation*

An associated periodicity value is calculated for each data trajectory to indicate the *quality* of the match to a model. The measure is a two-component periodicity function that checks how sinusoidal the data is (before the FFT fit) and how well aligned the start and end points of the position trajectory are for the cycle (should be the same for a closed pattern). If a translation exists in the model for the motion, the un-translated (and shifted) version is used for these calculations. The individual periodicity measure for each trajectory is a product of a discrete frequency magnitude ratio and a time-based tail alignment:

$$P = \left(1 - \frac{\max(|X[f \neq f_0]|)}{|X[f_0]|}\right) \times \left(1 - \frac{|x[t_a] - x[t_b]|}{|X[f_0]|}\right) \quad (5)$$

where  $f_0$  is the frequency containing the maximum energy,  $|X[f]|$  is the magnitude at frequency  $f$ , and  $(x[t_a], x[t_b])$  are the start and end point locations of the cycle. The time-based tail alignment checks to see how well the pattern is *closed*. The term  $\max(|X[f \neq f_0]|)$  is the maximum magnitude value in the discretely sampled frequency magnitude spectrum other than at the maximum  $|X[f_0]|$  (See Figure 23). In the target case of sinusoidal motions, this residual maximum should be small (even as a harmonic peak). If there is no motion (i.e.  $|X[f_0]|=0$ ), the resultant periodicity defaults to 1. The individual periodicity measures for each trajectory are multiplied together to yield a single periodicity measurement for the pattern. This measure is returned when a correct structural match is found at a velocity zero-crossing location to reflect the “goodness” of that match.



**Figure 23.** Frequency magnitudes for a single cycle trajectory. Components for the periodicity measure are noted.

#### 4.3.6 Recognition methods

Since we are interested in constructing an artificial perception system capable of recognizing and identifying categorical motions, we need to address how a recognition system might employ the resulting category models. For the simplest type of direct recognition of an unknown input pattern with the categorical models, we can perform two styles of recognition. The simplest approach to recognition in this case is a *goal-directed* scheme where a particular model of interest is matched to the unknown input. For example, a particular context or situation may suggest with high

probability that a certain motion will occur. Here, only a single model, or possibly a few models, will be actively used for recognition. If no contextual information were available to diminish the set of target models, we would alternatively perform full *category recognition*. This manner of recognition is an exhaustive search through all of the category models looking for a match. We will present recognition results using both of these approaches. Later in the conclusion, we speculate on other forms of recognition using the categorical models.

#### **4.4 Section summary**

In this section we described the computational implementation for recognition motivated by the categorical organization for the oscillatory motions. After describing related work in computer vision for recognizing periodic and oscillatory motions, we presented an experiment using infrared point-light arm motions that suggested the use of position trajectories, rather than velocity, as the underlying form of the data representation. We then described the details of the implementation method, focussing on lowpass filtering (Butterworth IIR), pattern segmentation (velocity zero-crossings), parameter estimation (FFT), computational organization (ordered motion sub-categories), and confidence measurement (periodicity calculation). We also briefly discussed two different methods for performing recognition of the input motion patterns to the models (goal-directed, full-category). We now examine this implementation with a range of experiments emphasizing the various aspects of the technique using synthetic motions, infrared point-light motions, and real video examples.



## 5 Experimental Results

To examine this categorical method for recognition and test the limits of the model, we first applied simulated data to see the effects of using non-ideal generating parameters. Due to the differing complexities of the motion models (residing at different computational levels), we next analyzed the model confusion for all the models and motions to report the nature and type of false matches. Then we tested the recognition system with human-generated point-light trajectories using the infrared tracking system. Lastly, to show that the system is applicable to real video motion, we examined the approach with video sequences of human, cartoon, and animal motions.

### 5.1 *Parameter estimation*

Here we examine the effects of using non-ideal motion generator values for the input signal when estimating the oscillatory parameters. Since we rely on the Fast Fourier Transform (FFT) to extract the sinusoidal parameters from each trajectory, we inspect here the quality of parameter estimation for synthetic non-ideal data using the FFT. The explicit parameter deviations determined from these experiments could be used as bounding thresholds for the parameters during recognition.

#### 5.1.1 *Number of data samples*

With a fixed sample rate of 30 Hz for the recognition system, a single cycle at different frequencies will have a different number of data samples per cycle. In this test, we examined the effect of this sampling of the generator function. To match the power-of-2 data length required for the FFT calculations, we linearly interpolated the data segments (originally from 15 to 120 samples per cycle) into either 64 or 128 samples. For these tests, single cosine waves of the form  $\cos(2\pi ft - \pi/2)$  were synthetically generated with a constant amplitude and phase, but with different frequencies ranging from 0.25 to 2.00 Hz. Table 4 shows that as the frequency is increased (resulting in a coarser sampling of the cosine function even after interpolation) there is an increasing degradation of the FFT estimated frequency values (and a slight degradation in the other parameters as well).

| <b>Table 4. FFT parameter calculation</b> |             |           |                 |             |
|---|-------------|-----------|-----------------|-------------|
| $f$ (c/s)                                 | Data length | Amplitude | Frequency (c/s) | Phase (rad) |
| <b>0.25</b>                               | 120         | 1.00      | 0.25            | -1.55       |
| <b>0.50</b>                               | 60          | 0.99      | 0.49            | -1.52       |
| <b>0.75</b>                               | 40          | 0.99      | 0.73            | -1.52       |
| <b>1.00</b>                               | 30          | 0.99      | 0.96            | -1.52       |
| <b>1.25</b>                               | 24          | 0.99      | 1.20            | -1.52       |
| <b>1.50</b>                               | 20          | 0.98      | 1.43            | -1.52       |
| <b>1.75</b>                               | 18          | 0.98      | 1.65            | -1.52       |
| <b>2.00</b>                               | 15          | 0.98      | 1.87            | -1.52       |

### 5.1.2 Gaussian noise

In a second test, we added Gaussian noise ( $\sigma=0.1$ ) to the synthetic data of frequency changes from the previous experiment to check the parameter estimation process for the motions in response to this type of noise. We note however that most noise of this nature will be removed by the lowpass filtering used in the actual recognition system. As shown in Table 5, the increasing degradation results for the frequency estimation remain similar to the previous results. But now the amplitude and phase responses have more fluctuation.

| <b>Table 5. FFT parameter calculation with noise (<math>\sigma=0.1</math>)</b> |             |           |                 |             |
|--|-------------|-----------|-----------------|-------------|
| $f$ (c/s)  | Data length | Amplitude | Frequency (c/s) | Phase (rad) |
| <b>0.25</b>  | 120         | 1.00      | 0.25            | -1.55       |
| <b>0.50</b>  | 60          | 0.99      | 0.49            | -1.53       |
| <b>0.75</b>  | 40          | 1.00      | 0.73            | -1.53       |
| <b>1.00</b>  | 30          | 0.99      | 0.97            | -1.52       |
| <b>1.25</b>  | 24          | 0.96      | 1.20            | -1.48       |
| <b>1.50</b>  | 20          | 1.02      | 1.43            | -1.54       |
| <b>1.75</b>  | 18          | 1.01      | 1.65            | -1.50       |
| <b>2.00</b>  | 15          | 0.93      | 1.87            | -1.53       |

### 5.1.3 Non-idealized frequency and phase

We now examine the calculated phase difference between two signals as affected by altering the frequency and phase in only one trajectory, simulating the generation of a motion pattern with non-idealized parameter values. Sets of sinusoidal waveforms were generated of the form:

$$\begin{aligned}x(t) &= \sin(2\pi f_x t + \phi_x) \\ y(t) &= \sin(2\pi f_y t + \phi_y)\end{aligned}$$

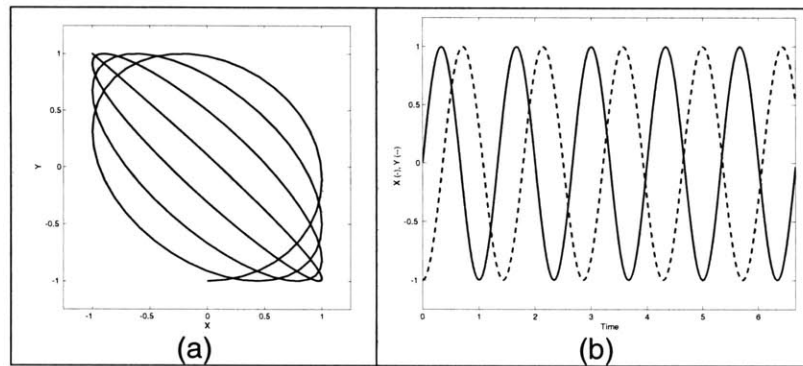
In the first test, a non-ideal circle motion was simulated by fixing  $f_x=0.75$ ,  $\phi_x=0$ , and  $\phi_y=-1.5708$ , but altering  $f_y$  from 0.5 to 1.0. Time sampling was again fixed to 0.0333 second intervals. After extracting a single pattern cycle (using three zero-crossings in  $\dot{x}(t)$  and starting at the first zero-crossing after  $t=0$ ), the FFT was applied to estimate the frequency and phase for each trajectory. All  $\hat{f}_x$  and  $\hat{f}_y$  estimations resulted in a frequency of 0.7324, except when  $f_y=0.50$  where no valid frequency other than the DC offset registered.

Results comparing the *calculated* phase difference from the *ideal-model* phase difference ( $\Delta\phi=(1.5708-(\hat{\phi}_x-\hat{\phi}_y))$ ) for the various  $f_y$  frequency values are shown in the first two columns of Table 6. The values in bold show permissible phase differences for recognition as governed by the lowpass filter threshold (0.3142 rad). This acceptance region (from the phase difference) for the  $f_y$  values other than 0.75 is short-lived, for as time progresses the two trajectories become periodically less synchronized (See Figure 24). Only the initial cycles of each synchronization are in fair alignment for recognition. Figure 25 shows patterns using a sample range of  $f_y$  values that yield acceptable phase differences.

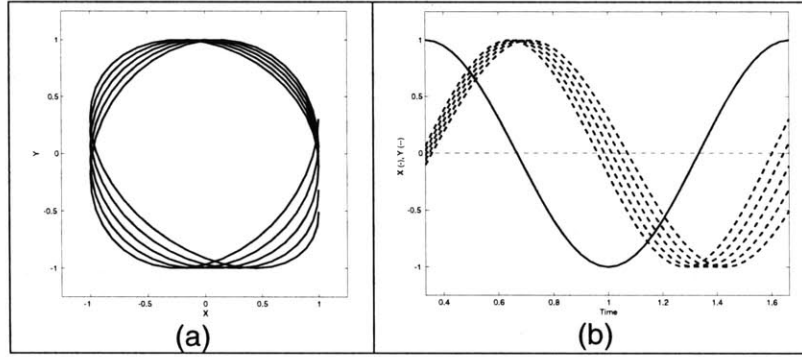
The second test fixed  $f_x=f_y=0.75$  and  $\phi_x=0$ , but systematically varied  $\phi_y$  from -2.3562 to -0.7854 ( $-\pi/2 \pm \pi/4$ ). Results for this test are shown in the second two columns of Table 6. As in the previous frequency-altering test,  $\hat{f}_x$  and  $\hat{f}_y$  were again calculated to be 0.7324. The permissible phase difference values (within tolerance from the lowpass filter error) from the model phase difference are shown in the table (in bold). Figure 26 shows patterns using a sample range of phase values that yield acceptable phase differences.

| <b>Table 6. Phase differences from model</b> |               |          |               |
|--|---------------|----------|---------------|
| $f_y$  | $\Delta\phi$  | $\phi_y$ | $\Delta\phi$  |
| 0.50   | 1.5264        | -2.3562  | 0.7829        |
| 0.55   | 1.1587        | -2.1991  | 0.6355        |
| 0.60   | 0.8400        | -2.0420  | 0.4774        |
| 0.65   | 0.5678        | -1.8850  | 0.3187        |
| 0.70   | <b>0.2981</b> | -1.7279  | <b>0.1595</b> |
| 0.75   | <b>0.0001</b> | -1.5708  | <b>0.0001</b> |
| 0.80   | 0.3376        | -1.4137  | <b>0.1593</b> |
| 0.85   | 0.6954        | -1.2566  | 0.3184        |
| 0.90   | 1.0325        | -1.0996  | 0.4772        |
| 0.95   | 1.3232        | -0.9425  | 0.6353        |
| 1.00   | 1.5708        | -0.7854  | 0.7927        |

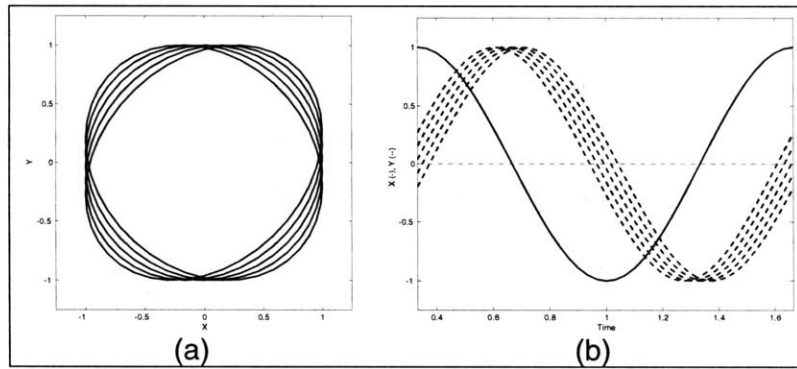
Phase difference results from model by altering frequency and phase components of  $x(t)=\sin(2\pi f_x t+\phi_x)$ ,  $y(t)=\sin(2\pi f_y t+\phi_y)$ . Default model values are  $f_x=f_y=0.75$ ,  $\phi_x=0$ , and  $\phi_y=-1.5708$ . Phase difference from model is given as  $\Delta\phi=(1.5708-(\hat{\phi}_x-\hat{\phi}_y))$ . Bold values show phase differences passing filter-based threshold ( $18^\circ=0.3142\text{rad}$ ).



**Figure 24.** Result of extended duration (5 cycles) with  $f_y=0.70$ . The spatial pattern quickly diverges from circular motion in (a). The temporal result with trajectories is shown in (b).



**Figure 25.** Range of circular motion accepted when  $f_y$  is altered. (a) Spatial patterns. (b) Temporal trajectories. Dashed lines show the valid range of trajectories.



**Figure 26.** Ranges of circular motion accepted when  $\phi_y$  phase altered. (a) Spatial patterns. (b) Temporal trajectories. Dashed lines show the valid range of trajectories.

## 5.2 Model confusion

In this categorical framework, the computations allocated to recognize the simpler motions are localized and unaware of the additional components (or specifications) that exist in the more complex motions (e.g. translation, frequency-doubling). Thus there may be unexpected results when models of one complexity are tested with motions of a different complexity. Here we examine confusion results of this type by testing each model with an example of each motion. This manner of testing using all the models for a particular input is similar to how *full categorical* recognition is performed. With our particular implementation and the relatively small number of models, it is not computationally taxing to examine the full set of category models for recognition.

The first test examining the model confusion employed synthetic motion trajectories. The data for this experiment were created by generating a point-light image sequence using computer graphics for each motion to simulate the video input of the infrared tracking system (See Figure 16). In addition, both performance directions (clockwise, counter-clockwise) were created for the appropriate motions to address their dual model phase values. The imagery was then processed by the recognition system as if the data were actually generated by the infrared tracking system.

For this test, the category models were each verified against all the synthetic data examples checking for multiple model responses to a single data trace. Only the structural components are of interest here and were examined for the matches. The stylistic information is not concerned with the pattern itself, and thus was ignored in this test. The results of the test are placed in the confusion matrix in Table 7. The correct matches for the models to the motions are located along the diagonal (in bold) in the table. Values shown are the maximum periodicity values (described in Section 4.3.5) returned when a structural match was found (out of multiple cycles at the velocity zero-crossings). The remaining values in the table represent the false matches. In these cases, the best-fit sinusoids to the data derived from Fourier analysis resulted in a correct structural match, though with a much smaller match score (periodicity measure).

As shown in the table, the circle model additionally matched the spiraling and looping motions (though having low periodicity measures). These matches are not unreasonable because the simpler circle model has no concept of translation, and the more complex spiral and loop motion are derived from circle motion with the addition of a translation (translatory circle motion). The

circle model therefore did the best it could to explain these related motions. The one-time matches triggered for figure-8 motion by the circle model and for circle motion by the spiral model were due to the effects of the initial lowpass filter transient on the signal. Thus we can discount these errors as any meaningful response. The response of the loop model during the U-shuttle motion was the result of a single shuttle swing closely matching the motion between successive vertical looping peaks, though having a larger translation. This is also due to the ability of recognition to occur at any zero-crossing in the signal (using data through 2 zero-crossings earlier to extract a single pattern cycle) rather than only at the end of the complete pattern. Since the data in the incorrect matches are not truly the correct sinusoidal fit to the models, the resultant periodicity measures are lower. Bold values in the table highlight the correct matches, which show that thresholding the periodicity match values at 0.8 yields correct discrimination (separation) for the models.

| <b>Table 7. Structural recognition confusion matrix for synthetic data.</b> |             |             |             |             |             |             |             |             |
|---|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Model   | Motion      |             |             |             |             |             |             |             |
|   | Up-down     | Left-right  | Circle      | Spiral      | Undulate    | Loop        | Figure-8    | U-shuttle   |
| Up-down   | <b>0.97</b> | -           | -           | -           | -           | -           | -           | -           |
| Left-right  | -           | <b>0.97</b> | -           | -           | -           | -           | -           | -           |
| Circle  | -           | -           | <b>0.95</b> | 0.14        | -           | 0.49        | 0.09        | -           |
| Spiral  | -           | -           | 0.17        | <b>0.92</b> | -           | -           | -           | -           |
| Undulate  | -           | -           | -           | -           | <b>0.97</b> | -           | -           | -           |
| Loop  | -           | -           | -           | -           | -           | <b>0.94</b> | -           | 0.77        |
| Figure-8  | -           | -           | -           | -           | -           | -           | <b>0.94</b> | -           |
| U-shuttle   | -           | -           | -           | -           | -           | -           | -           | <b>0.98</b> |

Entries report periodicity values of structural matches. Bold values indicate correct matches, and slots with - represent no structural match.

We next applied the system to actual user-generated data trajectories using the infrared tracking system. In this second test, we gathered example trajectories (each of a few cycles) for each of the oscillatory patterns from two individuals (subjects DH and BT). The results are shown in Table 8 and Table 9. The italic values in these tables show errors that were not present in the synthetic test. All of the correct matches were found with high confidence (periodicity measures) except for the spiral motion that was not identified in either of the users (phase errors above threshold). This pattern (See Figure 27) is difficult to *perform* by arm movements with a wand and difficult to *recognize* with 2-D trajectories (may require 3-D information for these actual motions). We kept the system based on the 2-D information because spiraling is the only 3-D pattern in the category and its projection still appears non-accidental, looking similar to “looping skyward”. If we open the phase threshold from 0.3142 (18 degrees) to 0.8 (46 degrees), then the system identifies the

spirals for subjects DH (max periodicity measure of 0.82) and BT (max periodicity of 0.50). With this new threshold no additional false alarms appeared in subject DH, and only a single match additionally appeared for BT with the U-shuttle (periodicity measure 0.07). Therefore the original threshold choice may be too limiting with the real motion data for this particular pattern.

Upon inspection of the remaining errors, we found that the circle model also recognized (with low score) the U-shuttle motion because the shuttle swings curved in at the peaks enough to almost produce a circle-like pattern. The low matches to the figure-8 were because some lobes barely appeared like a circle. The spiral model matched the circle, loop, and U-shuttle motions by identifying a slight translation. The loop model also had a one-time match for circle motion, which was found to have a small translation. We find that the user-generated motions contain mostly the same types of confusion errors as found with the synthetic data, but with additional confusion between the translatory patterns. Again, we generally see match values (maximum periodicity found in data input) with higher values for the correct matches (in bold) than for the errors.

If we apply a “winner-take-all” strategy to these point-light motions for subjects DH and BT, by picking the model with the highest periodicity score for an unknown input motion pattern, we can regularly select the correct recognition model. We see in these tables that the correct model for a motion usually has a much higher periodicity match score when compared with the other models (examining the values in a particular column of the confusion table). This strategy only applies when models signal a match at the same zero-crossing data locations. For example, the loop, figure-8, and U-shuttle models all look for matches at the velocity zero-crossings in the vertical y trajectory. But, the circle model looks for matches at the zero-crossings in the vertical y or horizontal x trajectory. Therefore this circle model can have a response (at the horizontal zero-crossings) when the other models are not looking for a match. Also the extent of the motion is important. We see that a single swing of the U-shuttle can appear as the motion between two vertical looping peaks. Thus only *three* vertical zero-crossings are needed for the loop model, but *five* vertical zero-crossings are required for the U-shuttle model.

Of the matches reported for subjects DH and BT, we find only a few cases where the winner-take-all approach failed to select the appropriate model. For subject DH, the loop and spiral model each had a one-time higher periodicity score (loop: 0.65, spiral: 0.20) slightly higher than the correct circle model (0.52, 0.13 respectively). For subject BT, the loop model had a few



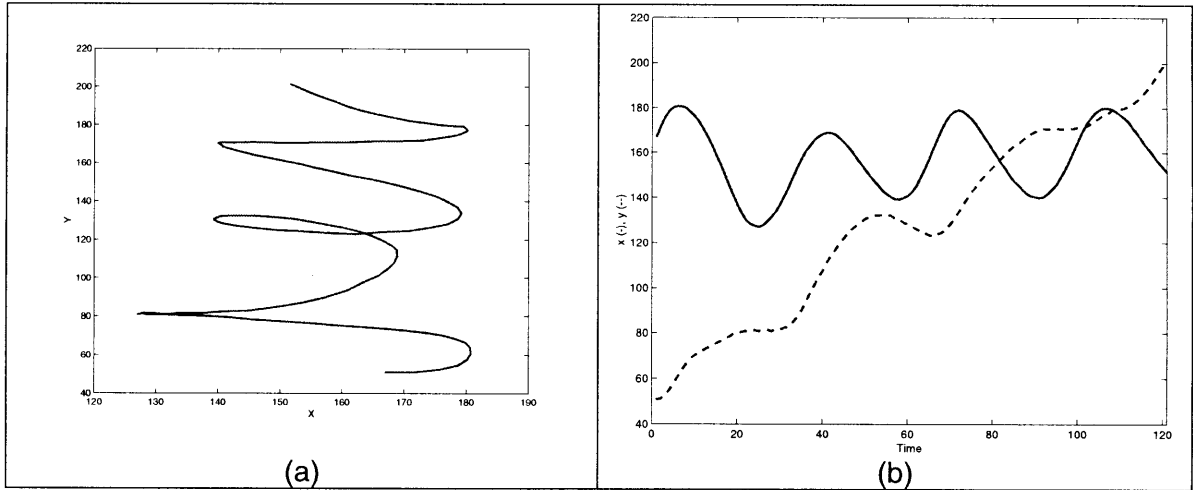
responses higher than the correct U-shuttle model (as described earlier, a *single* swing of the U-shuttle appears like the motion between two vertical looping peaks). Also, one match of the spiral model had a higher response (0.19) than the correct circle model (0.13), though both had small periodicity scores. But generally, the correct models respond with a higher periodicity measure for the appropriate motions, giving a winner-take-all strategy a high chance of success (though not guaranteed to be 100%, as shown above).

| <b>Table 8.</b> Structural recognition confusion matrix for subject DH. |             |             |             |        |             |             |             |             |
|---|-------------|-------------|-------------|--------|-------------|-------------|-------------|-------------|
| Model   | Motion      |             |             |        |             |             |             |             |
|   | Up-down     | Left-right  | Circle      | Spiral | Undulate    | Loop        | Figure-8    | U-shuttle   |
| Up-down   | <b>0.86</b> | -           | -           | -      | -           | -           | -           | -           |
| Left-right  | -           | <b>0.93</b> | -           | -      | -           | -           | -           | -           |
| Circle  | -           | -           | <b>0.80</b> | -      | -           | 0.57        | 0.16        | <i>0.11</i> |
| Spiral  | -           | -           | 0.20        | -      | -           | <i>0.39</i> | -           | <i>0.04</i> |
| Undulate  | -           | -           | -           | -      | <b>0.88</b> | -           | -           | -           |
| Loop  | -           | -           | <i>0.65</i> | -      | -           | <b>0.83</b> | -           | 0.54        |
| Figure-8  | -           | -           | -           | -      | -           | -           | <b>0.70</b> | -           |
| U-shuttle   | -           | -           | -           | -      | -           | -           | -           | <b>0.61</b> |

Entries report periodicity values of structural matches. Bold values indicate correct matches, italic values show errors not found in synthetic data, and slots with - represent no structural match.

| <b>Table 9.</b> Structural recognition confusion matrix for subject BT. |             |             |             |        |             |             |             |             |
|---|-------------|-------------|-------------|--------|-------------|-------------|-------------|-------------|
| Model   | Motion      |             |             |        |             |             |             |             |
|   | Up-down     | Left-right  | Circle      | Spiral | Undulate    | Loop        | Figure-8    | U-shuttle   |
| Up-down   | <b>0.91</b> | -           | -           | -      | -           | -           | -           | -           |
| Left-right  | -           | <b>0.96</b> | -           | -      | -           | -           | -           | -           |
| Circle  | -           | -           | <b>0.83</b> | 0.02   | -           | 0.59        | 0.02        | <i>0.13</i> |
| Spiral  | -           | -           | 0.19        | -      | -           | -           | -           | -           |
| Undulate  | -           | -           | -           | -      | <b>0.68</b> | -           | -           | -           |
| Loop  | -           | -           | -           | -      | -           | <b>0.90</b> | -           | 0.83        |
| Figure-8  | -           | -           | -           | -      | -           | -           | <b>0.74</b> | -           |
| U-shuttle   | -           | -           | -           | -      | -           | -           | -           | <b>0.71</b> |

Entries report periodicity values of structural matches. Bold values indicate correct matches, italic values show errors not found in synthetic data, and slots with - represent no structural match.

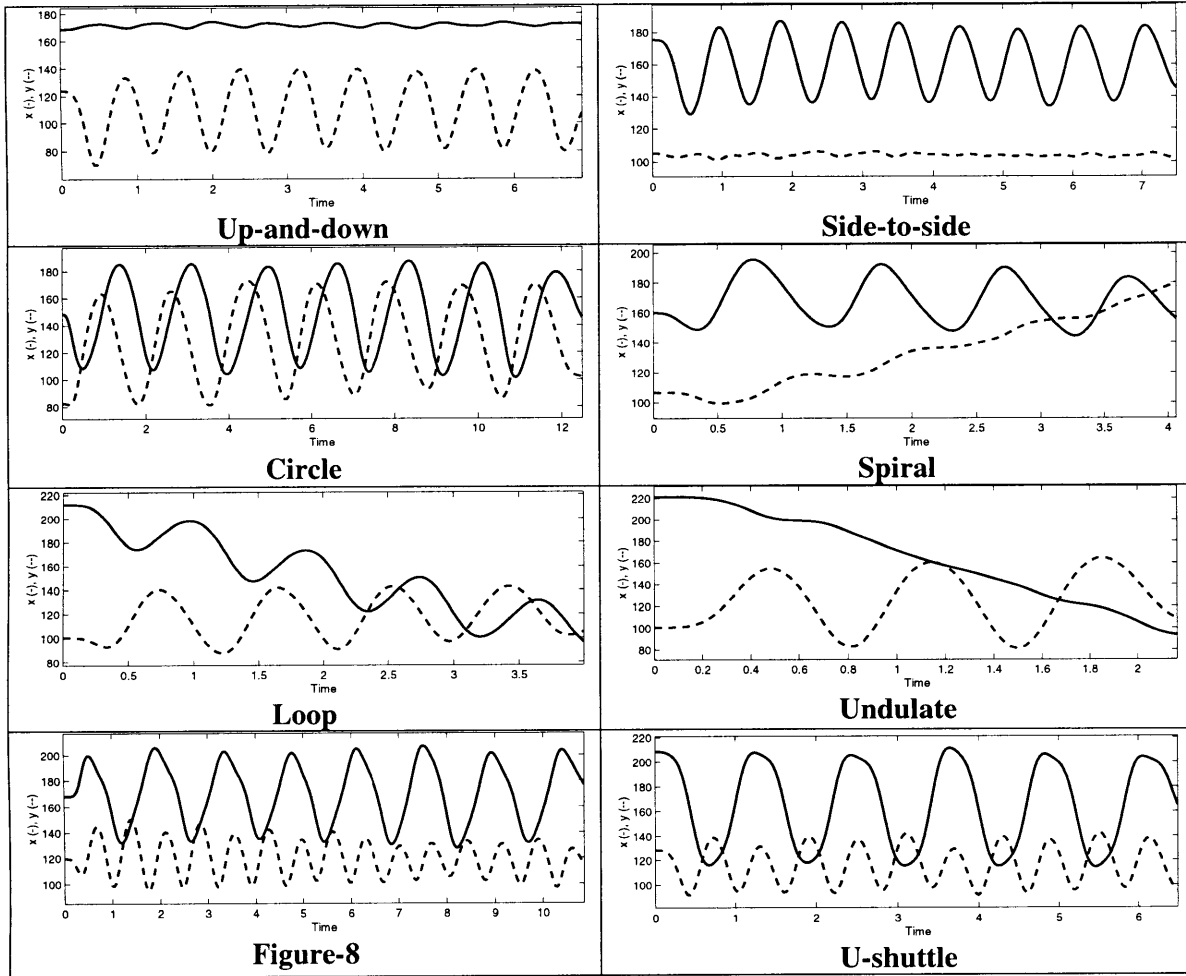


**Figure 27.** Spiral motion pattern of subject BT. (a) Spatial pattern in the image. (b) Position trajectories (x,y). This 3-D pattern was difficult to recognize as a spiraling motion in 2-D.

### 5.3 Goal-directed recognition

We further tested the recognition system to user-generated patterns for all of the motions with a third individual employing the infrared ball-tracking system. For these tests, the motions for each of the eight patterns were created by a single user with the tracking wand, and were digitized and processed offline. Then we performed *goal-directed* recognition by examining the result of applying the correct model to the input motions. The user-generated trajectory plots of the motions used for this test are shown in Figure 28.

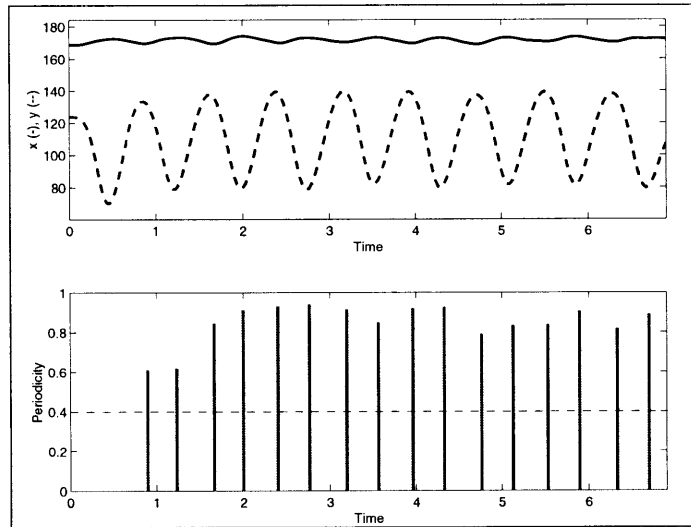
The recognition results using the structural information are shown in Figure 29 – Figure 36. The top plot of each figure shows the x, y position traces as solid and dashed lines, respectively. The bottom plot of each figure shows the structural recognition results. The spike magnitudes reflect the periodicity measures at the appropriate velocity zero-crossing locations in the above traces. Overall results show strong recognition throughout the user motions, though the periodicity values can be up to 50 percent lower than those recovered from the synthetic data in the previous experiments. The decrease is expected for *natural* motions. This may suggest a periodicity threshold of 0.4 for recognition, as shown in the bottom match response plots with a dashed horizontal line (possibly followed by the winner-take-all strategy if comparing to other models). The few low periodicity measures at the start of some motions are a result of the initial IIR filter transient.



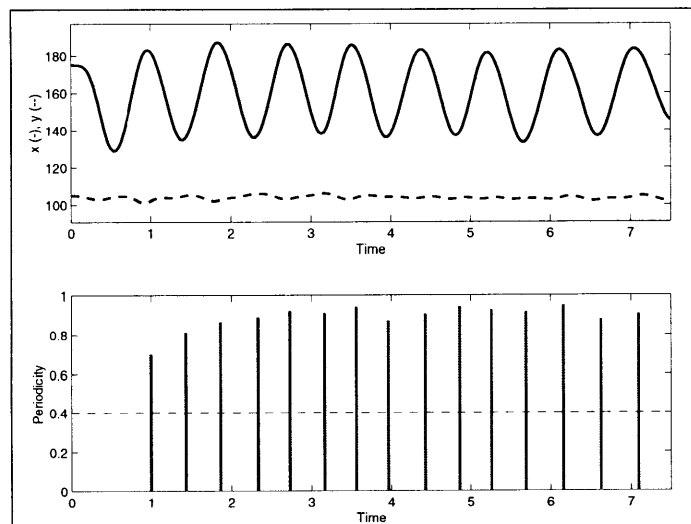
**Figure 28.** Motion trajectories for the oscillatory motion category generated by subject JD using the infrared tracking system. Solid and dashed trajectories correspond to the x and y trajectories, respectively.

As shown in the figures, the recognition system produced strong results to the user motions. A few missed targets were however found in the spiral and undulate motions (shown in Figure 32 and Figure 33, respectively). For the spiral motion, the first and last match error were due to a phase difference above tolerance and the remaining match errors were due to the translatory y-motion amplitude being smaller than the scale threshold (10%) in relation to the non-translatory x-motion (and thus was removed from consideration as being sinusoidal). As stated before, a likely improvement may involve using the x, y, and z motion trajectories for recognition, where the system would match the circular relations for x and z, and identify a translation in the y trajectory. Or we could increase the phase difference threshold. For the undulate motion, the match error at the beginning was caused by the amplitude of the translatory trajectory (x-motion)

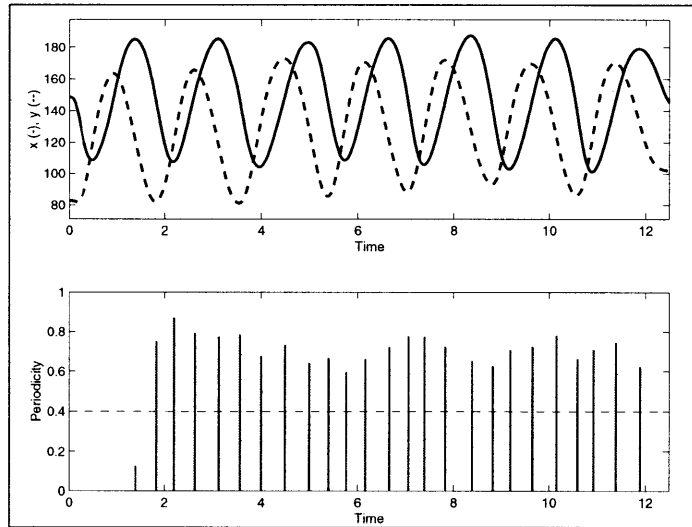
registering a visible non-zero amplitude (a bump) when it should have been a linear translation only. Thus this variation corrupted the structural match making the recognition system identify more than just a translation. The remaining models/motions show correct recognition.



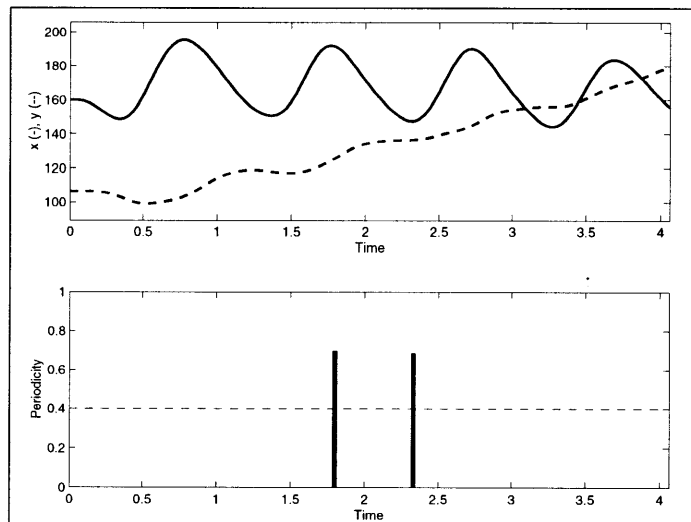
**Figure 29.** User position trajectories and matching results for **up-and-down** motion.



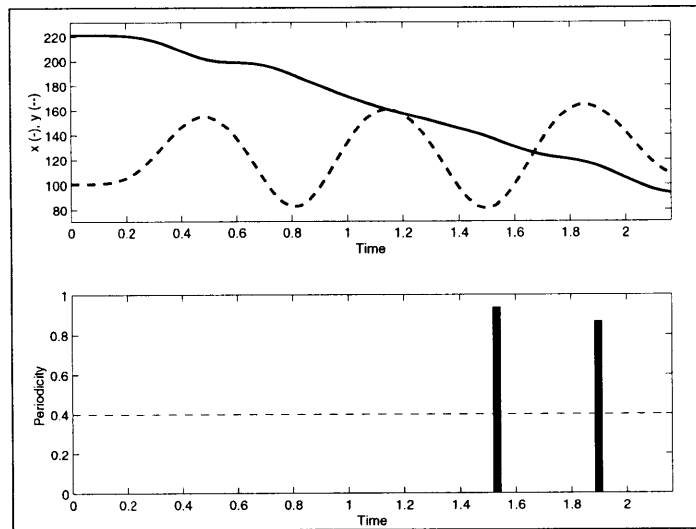
**Figure 30.** User position trajectories and matching results for **side-to-side** motion.



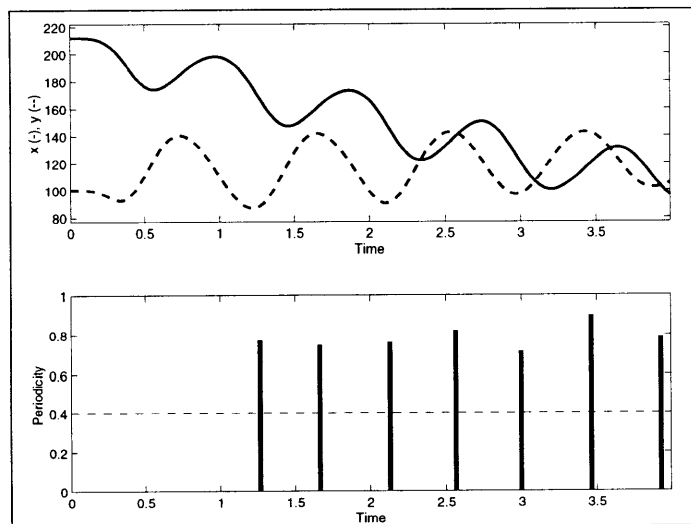
**Figure 31.** User position trajectories and matching results for **circle** motion.



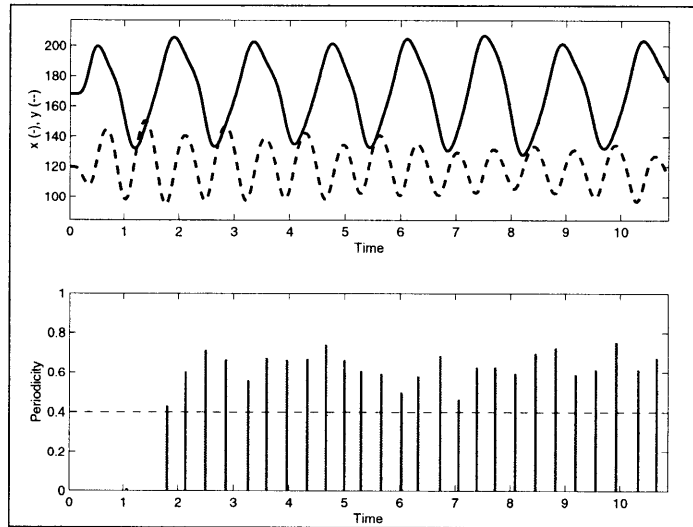
**Figure 32.** User position trajectories and matching results for **spiral** motion.



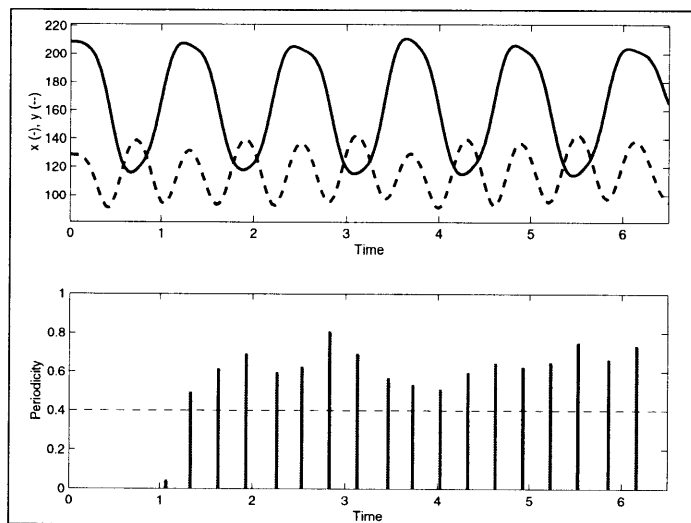
**Figure 33.** User position trajectories and matching results for **undulate** motion.



**Figure 34.** User position trajectories and matching results for **loop** motion.



**Figure 35.** User position trajectories and matching results for **figure-8** motion.



**Figure 36.** User position trajectories and matching results for **U-shuttle** motion.

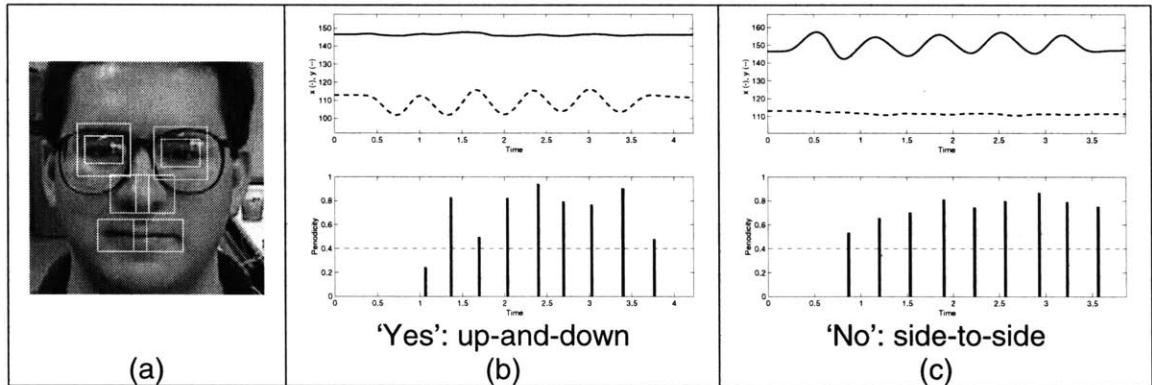


## 5.4 Real video

To show the extendibility of the system beyond synthetic and infrared point-light display input, we lastly demonstrate the recognition system using real video sequences of oscillatory motions. We begin by examining an automatically tracked person's face making simple side-to-side and up-and-down gestures. Then we show the system with a segment of classic animation containing a looping motion. Lastly, the recognition system is applied to a video sequence of a gibbon swinging across a rope. In this last example, a new swinging model is derived as an extension to the original category of motion.

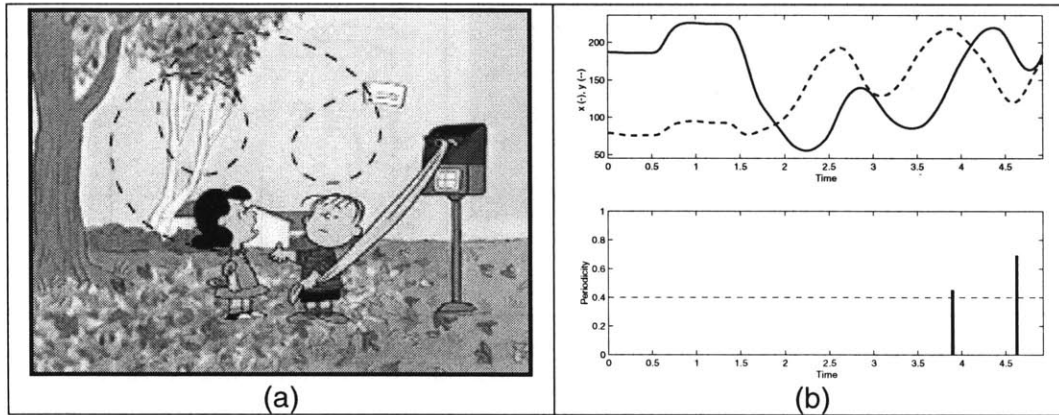
*"...if we wish to indicate 'Yes' or No' by movements of the head, there is little more we can do, except by great effort, than turn the head from side to side or wag it backwards and forwards. ...All that remains is to accept as a convention that nodding or shaking the head shall mean 'Yes' or 'No'." [20]*

As described above, one of the simplest and most basic of human gestures is nodding or wagging the head to indicate a 'Yes' or 'No' response, though the meaning of the gestures may be culturally specific. To recognize these head movements, we applied an automatic face detector and real-time face tracker [53] to extract the motion of the head performing the gestures. The face detector uses skin classification, symmetry operations, and eigenfaces to identify a face. The face tracker then utilizes structure-from-motion to generate a 3-D model of the face from Kalman-filtered trajectories of 2-D image patch locations and parameterized models of facial structure. The region-based face model used for tracking is marked on the face image shown in Figure 37.a. The output of the face tracker (and the input to the recognition system) is the location of the middle of the face at the nose location. The structural recognition matches for the up-and-down ('Yes') and side-to-side ('No') gestures of the head are shown in Figure 37.b,c. These results show correct recognition of the motions. We also tested these simple motions against all of the oscillatory models with the result of only the correct 1-D model signaling a match to the motions. As verbal and head gestures naturally occur together during communication, recognition of such common head gestures, coupled with speech analysis systems, could be exploited as a useful human-computer interface. The stylistic frequency value for the oscillations could additionally be used to signify the intensity of the expression.



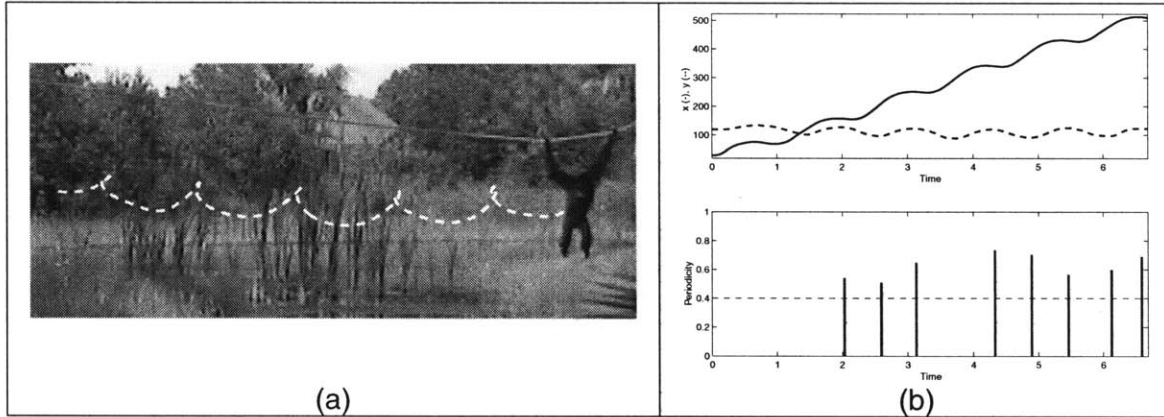
**Figure 37.** Structural recognition of YES and NO head gestures. (a) Face model overlaid on the input image. (b) Recognition results for up-and-down YES gesture. (c) Recognition results for side-to-side NO gesture. No other models reported a match to these motions. The x and y position trajectories are shown as solid and dashed lines, respectively. The spikes in the bottom plots show the quality of recognition at the sinusoidal velocity zero-crossings.

In addition to their role in biological motion, we can also find the oscillatory motions being used in other domains as interesting motion patterns. For example, the cel animation shown Figure 38.a depicts a letter being *looped* overhead of the characters into a mailbox (the dashed trajectory in the image was added to show the motion path of the letter over the entire scene). We manually extracted the position of the letter in each frame of the scene and analyzed its motion as an overhead looping pattern (clockwise circle motion with a horizontal translation). As shown in Figure 38.b, the system correctly identified the completion of the first loop and its continuation into the first half of the second loop at the velocity zero-crossings. There was no recognition for the end of the second loop because the letter did not complete the final loop in the scene (in the lowpass filtered data). But since we look for matches at the zero-crossing locations, we were able to identify the first half of the last loop. When this motion was tested with the other models, only the circle model registered a match (maximum periodicity score of 0.32).



**Figure 38.** Cel animation showing a letter being *looped* into the mailbox. (a) The dashed motion trace was added to the original animation to illustrate here in a single frame the movement of the letter. (b) Recognition results of the looping motion. (Video selection from “It’s the Great Pumpkin, Charlie Brown”, 1966. With permission from Lee Mendelson Film Productions, Inc. and United Media.)

One of the main advantages to using a category as the underlying representation for the motions is that it becomes easy to extend the category with additional motion models. Because of a single characterizing model (e.g. sinusoids), new models are created within the category by simply altering the existing structural parameters, possibly obeying certain rules or constraints of the other motions. This is different in the training-based approaches where an additional set of examples is required to construct a new model. As previously mentioned, the translation in the looping motion must be less than the product of its corresponding amplitude and frequency ( $2\pi f \alpha_x$ ), otherwise it is stretched to become a swinging pattern. Due to the categorical nature of the motions, the only difference between swinging and looping is this translation boundary. We can therefore construct a new model for the swinging pattern by simply altering the translation toward this boundary for the existing looping model. The true mathematical boundary is at the product of the amplitude and frequency. Because of the lowpass filtering and appearance of the swings, we chose a lower threshold of fifty percent of the amplitude-frequency product to signify a swinging (and not looping) pattern. The earlier infrared tracking of subjects DH and BT had *looping* translations well below fifty percent.

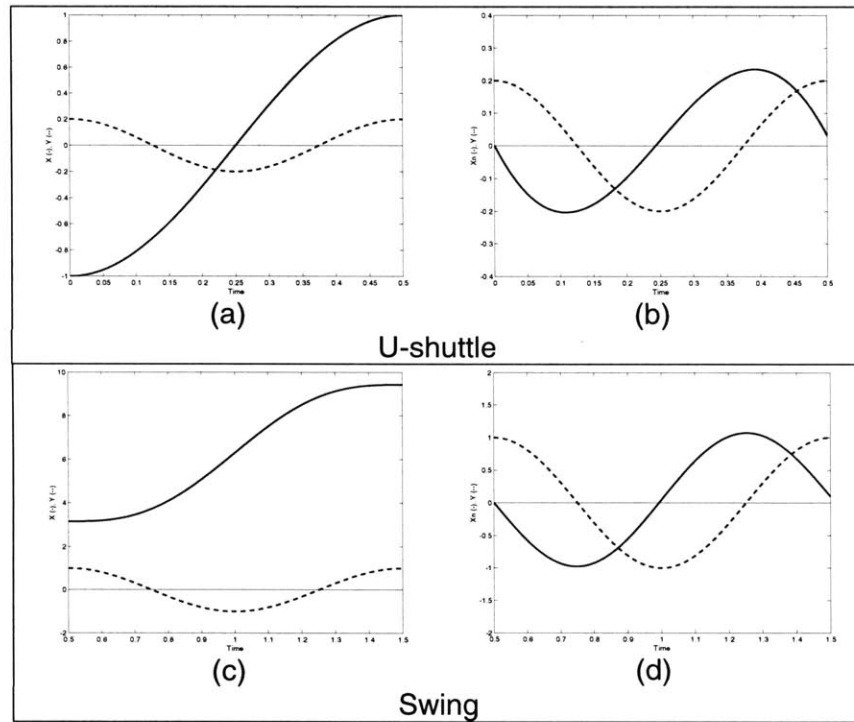


**Figure 39.** Structural recognition of the swinging pattern of a gibbon. (a) Gibbon sequence. (b) Recognition results. (Filmed by the author with permission at the Brevard Zoo, FL.)

We applied the newly derived swinging model to a video sequence of a gibbon swinging (brachiating) across a rope. The gibbon trajectory was automatically extracted and tracked in the sequence. The body of the gibbon was segmented by reverse-thresholding the intensity images (setting a pixel if the intensity is below a brightness value and discarding it if it is above that value) and collecting the centroid locations throughout the sequence of the dark region corresponding to the gibbon. Figure 39.a shows this trajectory overlaid on the final frame of the sequence. Results using the new swinging model from the category are shown in Figure 39.b. All the swings were correctly identified except for one portion of one swing. This singular error was due to a phase difference above threshold tolerance. When all the remaining models were tested with this sequence, only the circle model had a response (maximum periodicity response of 0.05). Most of the gibbon swings had a translation above eighty percent of the frequency-amplitude product. The immediate success of this model illustrates the power and extendibility of the categorical approach.

An interesting equivalence between this swinging model and the U-shuttle pattern offers a secondary approach for the recognition of pendular U-shuttle movement. Verbally described as “swinging back-and-forth”, U-shuttle motion has the repeated appearance of swinging-left and swinging-right along the same U-shaped path. It can be shown that the trajectories for each swing of the U-shuttle (leftward or rightward) are equivalent to an actual swing motion when *applied* to the swinging model. As shown in Figure 40, when a supposed translation is removed from a half-cycle of a U-shuttle motion (rightward swing) followed by a shift (as performed for motions like swinging and looping in the translation sub-category for recognition), the resulting trajectories

(See Figure 40.b) are found to be structurally equivalent to a corresponding translation removal and shift in a swing motion pattern (See Figure 40.d). Even though the full pattern cycle for U-shuttle has no translation, its half-cycle can be interpreted as being a sinusoid with a translation component. The significance of this equivalence is that we can construct an alternate method to recognize U-shuttle motion using a dual-state swinging mechanism where the states match to swing-left and swing-right patterns (instead of using the frequency-doubling computations).



**Figure 40.** Equivalence between U-shuttle half-cycle and swing motion. (a) Half-cycle trajectories of U-shuttle. (b) Trajectories after a translation removal and shift is applied to the x trace. (c) One cycle of swing. (d) Trajectories after translation removed and shift applied to x trace. Plots (b) and (d) result in equivalent motion patterns. The x and y motions are shown as solid and dashed lines, respectively.

## ***5.5 Section summary***

The experiments presented and discussed in this section examined the quantifiable results of the categorical implementation. The results of the method with synthetic, point-light, and real video motions demonstrated that such a simple (and more general) categorical implementation can capture the form of the motions when the computations are categorically organized. An additional advantage of the categorical approach was demonstrated by extending the original category to include a swinging pattern by simply changing a single parameter value of an existing looping model. The experiments in this section were performed in isolation to report and quantify the results. In the next section, we present a more applicable use of the models by connecting the resulting recognition of the motions to behavioral reactions of artificial creatures within an interactive artificial life environment.

## 6 Interactive Artificial Life

To further explore the recognition of the category models, we look to integration of the approach into interactive artificial life environments. With ethological aspirations, creature perceptions and actions in the artificial world are tightly coupled to produce rich interactions with other artificial creatures or with humans exchanging intentions with them. With relation to our interests, we can employ the categorical recognition approach for one creature to identify the corresponding intentional motion displays of another creature (artificial or human). Recognition then would be used to affect the behavioral response of the watching creature. Here the notions of success and failure of the models are determined in part by the resulting behavioral responses of the creatures based on the information derived from the visual input. Thus the results of recognition are directly observed in the reactive behaviors of the creatures.

As a demonstration of these concepts, we designed and constructed an interactive virtual world populated by artificial hummingbirds. The interactions in this world are facilitated by a user who controls the motions of a male hummingbird that is trying to convey his intentions toward a watching female hummingbird. The output of the recognition system watching the user is connected to the female hummingbird, where the recognition results are directly mapped into the female's behavioral motivations. Thus a real-time behavioral interaction between the user-controlled male hummingbird and reactive female ensues. Before describing the details of the system, we briefly review the four most influential artificial life and interactive systems that affected the design and goals of this virtual hummingbird world.

### 6.1 *Artificial life environments*

In this section we highlight those artificial and interactive environments that motivated the interactive virtual hummingbird world. The combinatory result from these inspiring systems was to have autonomous creature behavioral responses ([12, 98]) related to interactions performed in the real world ([14, 31]).

In Terzopoulos' artificial fishes scenario [98], a computational framework for creating fully functional artificial animals (fish) was given (See Figure 41.a). Artificial fishes in this domain are

autonomous agents with functional bodies controlled by artificial brains with synthetic visual perception. Here the animal's perceptual awareness of its world and behavior are effectively modeled. Terzopoulos' work also implements primitive reflexive behaviors that directly couple the creature's perceptions and actions. Primitive behaviors are combined into motivational behaviors whose activation depends on the artificial fish's mental state. These artificial fish have multiple behavior routines: avoiding-static-obstacle, avoiding-fish, eating-food, mating, leaving, wandering, escaping, and schooling. The long-term goal of that research is a computational theory that can potentially account for the interplay of physics, locomotion, perception, behavior, and learning in higher animals.

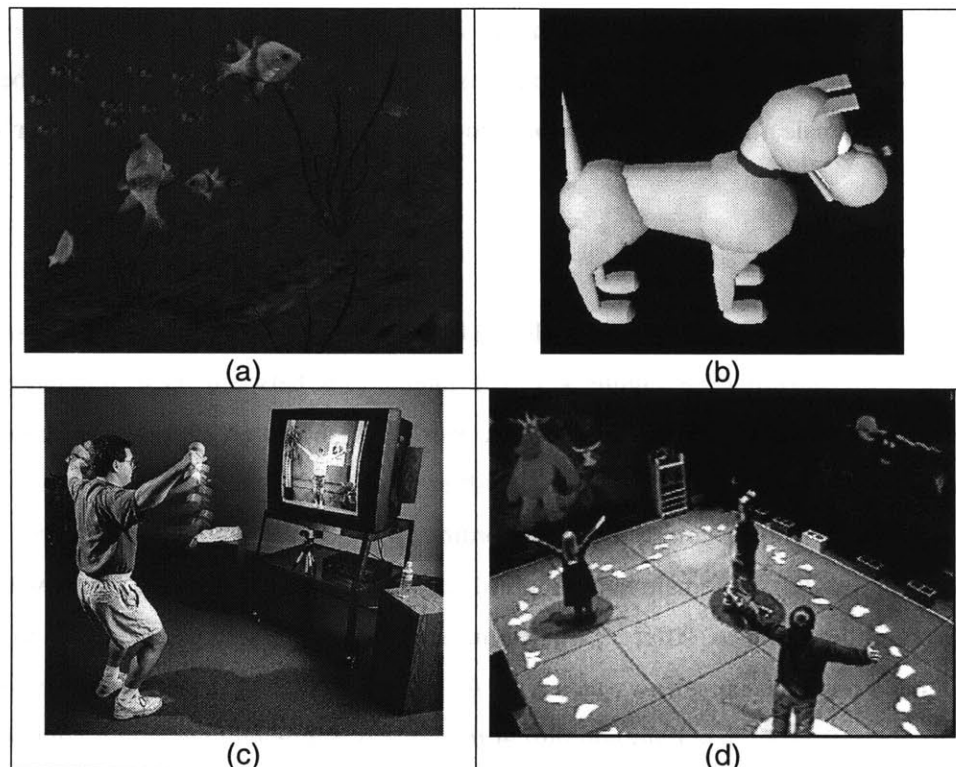
In Blumberg's research with Silas T. Dog – an interactive artificial dog (See Figure 41.b) – the focus was on building autonomous animated creatures which live in virtual 3-D environments and interact with people viewing the virtual world [12]. As with Terzopoulos' fish, the ethologically-inspired virtual dog in Blumberg's research was autonomous, in that at every instant it decided what to do based on what it could sense of the world, the state of its internal goals and motivations, and the actions which it could perform. A synthetic vision system (to perceive and interact within the virtual world) and camera-based vision system [110] (to interpret the gestures of people) were employed for Silas. A range of disciplines were exploited for this work and Terzopoulos', including artificial life, ethology, cognitive science, classical animation, computer graphics, and artificial intelligence. Both of these systems (artificial fishes, virtual dog) motivate the use of artificial creatures with internal motivations based on events perceived in their environment or the real world.

The interactive virtual aerobics instructor of Davis [31] was designed to interact and motivate a user as they together go through an aerobic workout (See Figure 41.c). The virtual instructor "watches" the user by means of a real-time computer vision system designed to recognize and interpret the exercise movements of the user [30]. The output of the recognition system (i.e. the activity of the user) and the responses (or state) of the virtual instructor were coupled in a simple reactive system to give the sensation that the instructor indeed watched the user. For example, if the person stopped performing the current exercise, the virtual instructor quickly responded "Get moving!". Or if the user was correctly exercising with the virtual instructor, he periodically encouraged the user saying "Good job!". The recognition of the exercise movements (or if not exercising) acted as the means of connection from the virtual instructor to the person. In this



system, the visually-based interaction focused on the activity of the user rather than on other virtual characters.

The KidsRoom [14] was created to be an interactive, narrative playspace for children using computer vision to facilitate the interactions. The experience attempted to create a fantasy-world within a child's bedroom where the children could participate in a storybook narrative and interact with virtual monsters that are projected onto the walls of the room (See Figure 41.d). The environment incorporated computer vision systems (e.g. [30]) to enable the virtual characters to interact with and respond to the children. For example, during one scene the children and virtual monsters had the opportunity to dance together, taking turns and copying the dance movements of each other. Motion recognition was used here to connect the actions of the children and the virtual monsters. As with the virtual aerobics instructor, this system motivated a focus of interaction that took place between the real and virtual world using the movements of the user(s) to drive the interaction.



**Figure 41.** Artificial life and interactive systems. (a) Artificial fishes. (b) Silas T. Dog. (c) Interactive virtual aerobics trainer. (d) The KidsRoom.

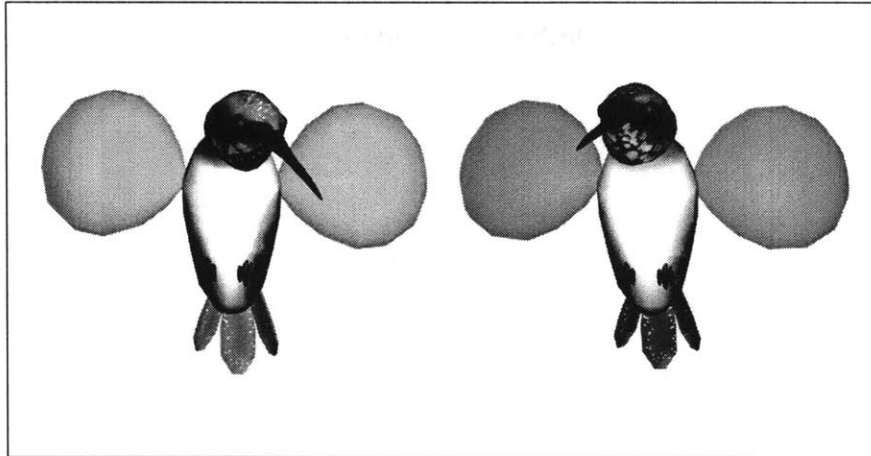
## 6.2 *Artificial hummingbirds*

Motivated by the artificial creatures, behavioral motivations, and interactive systems previously described, we designed an interactive virtual world with artificial creatures having behavioral responses that are affected by oscillatory motions performed by a person in the real world. Our 3-D virtual world is inhabited by artificial hummingbirds interacting around a feeder. Hummingbird creatures were chosen for the virtual world because of the rich behavioral responses using oscillatory motions that actually occur for the species. Also, designing and animating simple non-articulated hummingbird models makes lifelike believability simpler to attain without the need of complex motor skills for the creatures.

The basic scenario for this system involves a male hummingbird that is trying to signal his intentions to a watching female. As we are interested in perceptual recognition models of the categorical motions, this system relies on the recognition of particular oscillatory motions of the male hummingbird to affect the reactive behavior of the female. The infrared tracking system described earlier is employed to let a human control the motions of the male hummingbird. Recognition of the motions from the tracking system is then wired directly into the behavior system of the female, triggering behaviors from identification of the male's (user's) display motions.

### 6.2.1 *Virtual world*

Inhabiting this virtual world are a male and female hummingbird of the virtual species *Medialabus experimentii* (See Figure 42). The interactions between these hummingbirds take place around a centrally located feeder in the world, as shown in Figure 43. Throughout the experience, ambient hummingbird sounds (low-frequency buzzing/humming with high-frequency chirps) are played from a compact disk recording of Ruby-throated hummingbirds [35]. The experience begins with the female hummingbird contently flitting around the feeder with the male hovering off in the distance. After a brief moment, the male hummingbird flies up to the feeder and the female positions herself to watch him. At this point, the 2-D position control of the male is given to the user. The male hummingbird at this time is a directable creature, controlled by the user with the hand-held wand and infrared tracking system (as previously described in Section 4.2.1, and shown in Figure 44). The arm motions of the user (controlling the male hummingbird), as extracted by the infrared tracking apparatus, are analyzed by the vision system and the recognition results are sent to the behavior system to affect the interaction of the hummingbirds.



**Figure 42.** Virtual hummingbirds of the artificial species *Medialabus experimentii* (left: female, right: male).

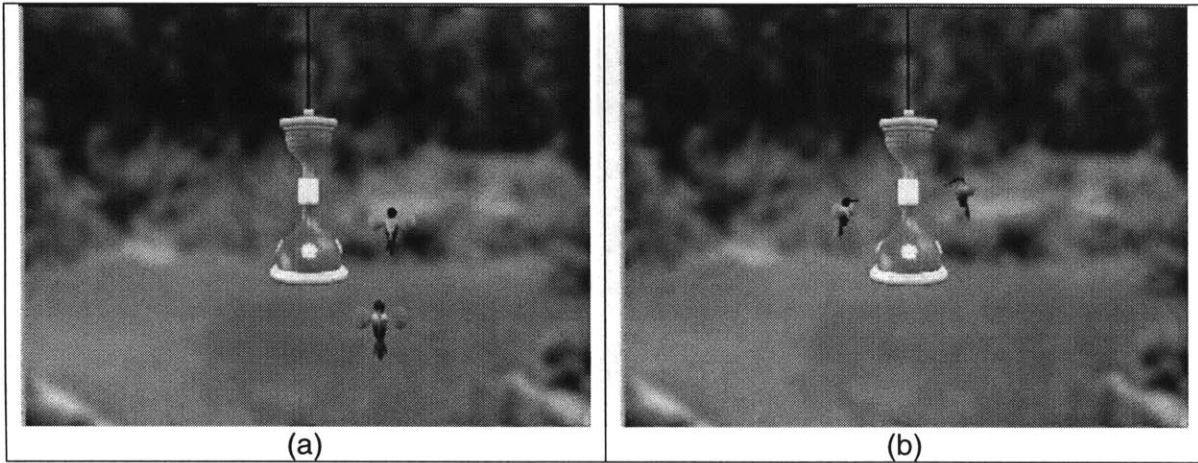


**Figure 43.** Virtual hummingbird world, with the male hummingbird on the right side of the feeder and the female on the left side.

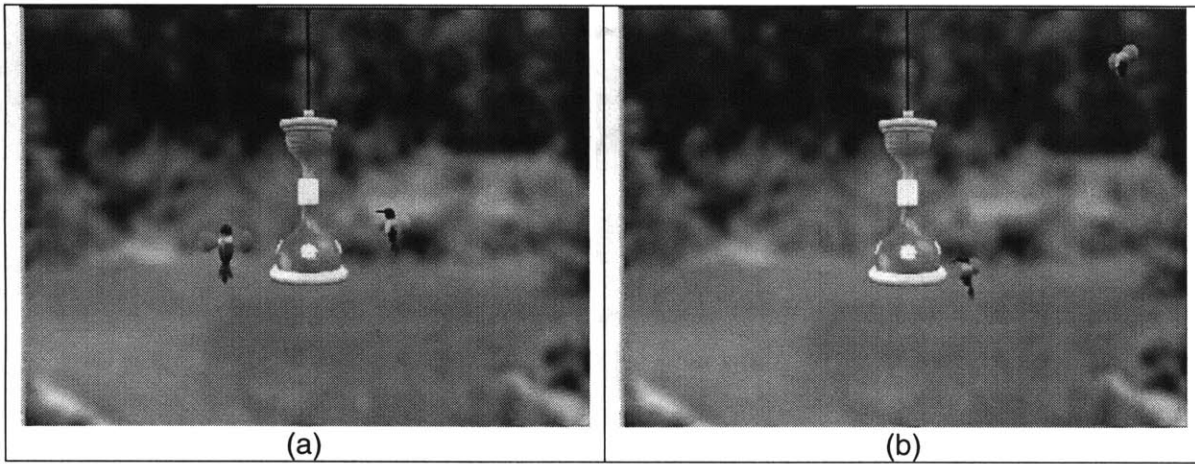


**Figure 44.** User interaction. The infrared tracking system enables the user to control with a wand the motions of the male hummingbird.

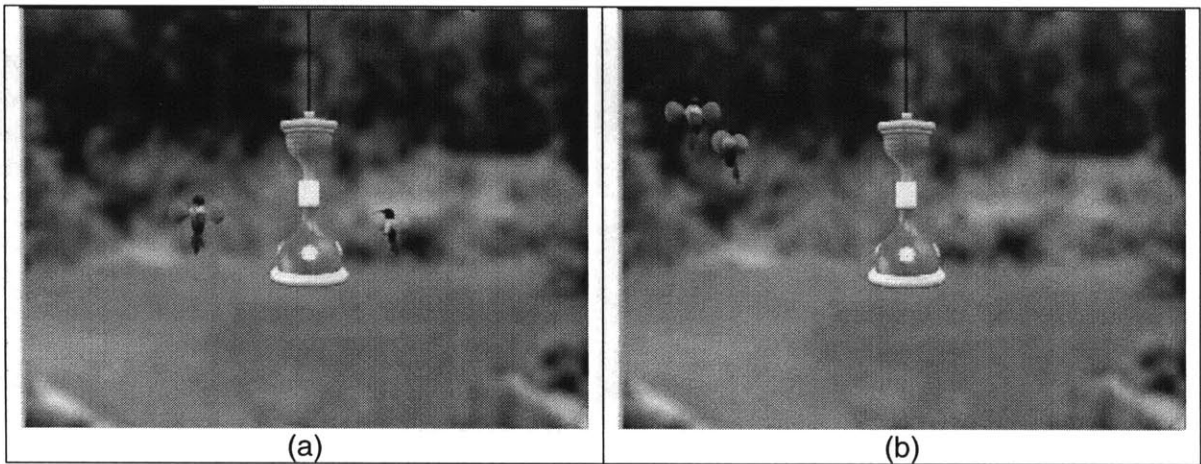
The person controlling the male hummingbird has the option of performing one of four oscillatory display motions that are directed at the female residing near the feeder. Recognition of these motions by the vision system then triggers a particular response in the female. The choice of movement displays for the user include repeated patterns of *up-and-down*, *side-to-side*, *figure-8*, or *U-shuttle* motions. This subset of the full oscillatory motion category contains two of the simplest (*up-and-down*, *side-to-side*) and two of the most complex (*figure-8*, *U-shuttle*) motions. These motions are used to affect the behavioral state, and thus the resulting reaction, of the watching female. Both the 1-D *up-and-down* and *side-to-side* motions of the male act as gestures to ask the female if he may come and join her at the feeder. Recognition of either of these two motion patterns by the category models in the vision system triggers an *acceptance* reaction by the female, where she returns to flitting around the feeder while permitting the male to join her (See Figure 45). A *figure-8* motion performed by the male is used as an aggression display to the female, attempting to bully her away from the feeder. Identification of this motion pattern causes *fear* in the female and she correspondingly backs away from the feeder leaving it to the pushy male (See Figure 46). Lastly, a *U-shuttle* movement of the male acts as a *courtship* signal for his intention of mating. Recognition of this motion results in the female joining the male and happily flying away with him (See Figure 47). If the user does not perform any of the gestures within a certain amount of time or does not perform any of the motions correctly (e.g. performs them sloppily), the female will not understand his intentions and will respond by aggressively defending the feeder and attacking him (See Figure 48).



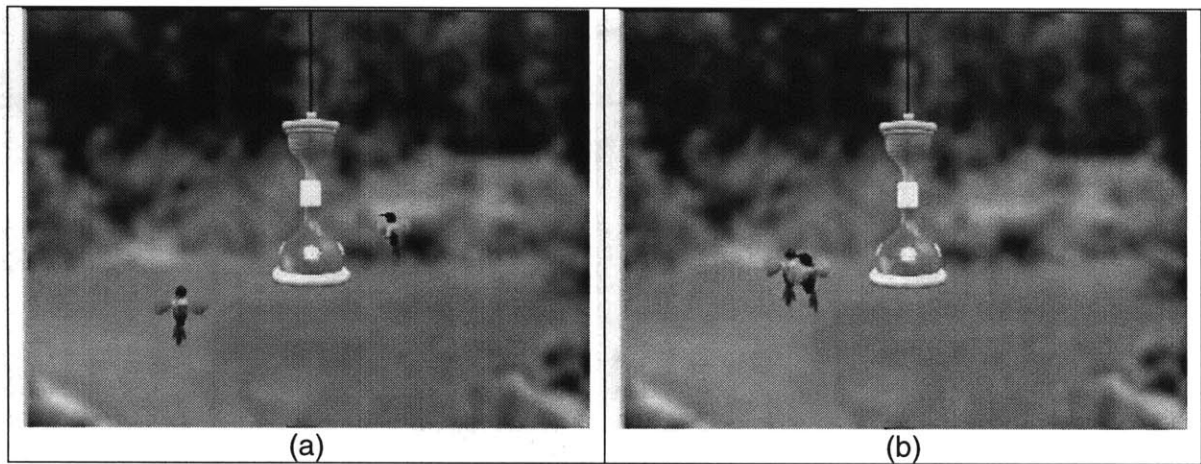
**Figure 45.** Acceptance behavioral response of female. (a) Male hummingbird performs up-and-down or side-to-side motion patterns. (b) Female accepts male.



**Figure 46.** Fear behavioral response of female. (a) Male hummingbird performs figure-8 motion pattern. (b) Female retreats, leaving the feeder to the male.



**Figure 47.** Courtship behavioral response of female. (a) Male hummingbird performs U-shuttle motion pattern. (b) Female and male fly happily away together.



**Figure 48.** Aggression behavioral response of female. (a) Male hummingbird does not perform any of the characteristic motion patterns. (b) Female aggressively defends the feeder by attacking the male.



### 6.2.2 Implementation

The vision system driving the interactions employs the category models described throughout this research. Only the computational levels/sub-categories for 1-D motion (up-and-down, side-to-side) and 2-D frequency-doubling motion (U-shuttle, figure-8) were active during recognition. The remaining computational levels of 2-D motion (circle) and translation (spiral, loop, undulate) were ignored for this demonstration. Recognition here can therefore be considered a limited form of *category* recognition or an extended version of *goal-directed* recognition. The thresholds for the system remained the same as were used for the previous experiments quantifying the results of the models, but with the addition of a match threshold at the velocity zero-crossings for the periodicity to above 0.4 (maximum of 1.0) to discard any uncharacteristic motions. The vision system had a constant rate of 30 Hz on Silicon Graphics O2 (R10000, IRIX 6.5). The system output at each frame to the behavior system was the tracking data to control the position of the male hummingbird (x and y position of the infrared ball in image coordinates) and the corresponding recognition results for the motion patterns (booleans of whether or not the system just recognized *up-and-down*, *side-to-side*, *figure-8*, or *U-shuttle*).

The behavior, motor, and graphics sub-systems for the virtual world were custom designed by the Synthetic Characters Group at the MIT Media Lab. Their Java-based implementation was written to simplify creature development by shielding the behavior and motor system from the details of the underlying graphics system. We describe here only the details of their system that are relevant to understanding the overall demonstration. See [56, 58, 87] for additional information.

The behavior system is responsible for selecting the appropriate actions for the hummingbirds. In our design we have a set of three primary motivational states that can be altered in the female: acceptance, fear, and courtship. These motivational states are affected by the on-line recognition results of the vision system tuned for the particular motion patterns that correspond to these states (*up-and-down*, *side-to-side*: acceptance, *figure-8*: fear, *U-shuttle*: courtship). A motivational state is defined as an accumulator with inputs, gains, and decay parameters. The general form of the motivational unit is

$$M_t = \alpha M_{t-1} + \beta I$$

where  $M_t$  is the motivational value at time  $t$ ,  $M_{t-1}$  is the previous motivational value with a fractional decay factor  $\alpha$  ( $<1.0$ ), and  $I$  is the input value (0 or 1) with gain factor  $\beta$ . The gain factor is used to increase the motivational value when a recognition match is found and the decay factor gradually returns it (non-linearly) to some base value when no matches are found. Therefore, the more often the pattern is seen in a particular time window, the higher the response will be. With this simple accumulate-decay model, a faster user performance of the motions and recognition (if no “stylistic” speed test is required) corresponds to achieving the motivational threshold faster, thus reacting similar to a supernormal stimulus for the motion pattern. These parameters can be individually set depending on the motion to recognize or the desired responsiveness of the creature. A resulting behavior associated with the motion pattern is triggered when the motivational value reaches a predefined threshold. This model ensures that single, random recognition matches do not prematurely trigger the reactions.

The three motivational units corresponding to the reactive behaviors (acceptance, fear, courtship) are run simultaneously keeping all motivational accumulators running concurrently. When one of the motivations is triggered above threshold, functional animations for both hummingbirds act out the resulting behavioral interaction (at this point the user control is ignored). The aggressive behavior by the female when no motions are recognized is activated only when none of the motivational units respond in the specified time for the interaction. The behavior system controlling the actions of the hummingbirds was run on a single processor Pentium III PC (500 MHz, Windows NT). Communications between the vision system (Silicon Graphics O2) and the behavior engine (PC) were achieved through a serial port connection (O2 to PC). Due to the processor demanding graphics, the rendering was piped from the PC running the behavior engine to a secondary machine (described below). With this system there is a small but noticeable delay between the user performance and corresponding motion of the male hummingbird.

The graphics system is responsible for rendering and managing the creature’s geometry. The hummingbird and feeder models were created in 3D Studio Max and exported to a VRML-based format. The non-articulated hummingbird and feeder models were custom designed with 2,900 and 10,000 polygons, respectively. Only the hummingbird models were texture-mapped for their appearance. The background image in the virtual world was created by blurring a scanned picture of an outdoor scene (to give it the appearance of being in the distance) and texture-mapping it onto a vertical plane (created in 3D Studio Max) placed behind the feeder. All movements of the hummingbirds are created either from direct user control (for the male hummingbird dances) or



by functional motor commands (e.g. “MOVE-TO-FEEDER”). No stored animation files were used in this implementation. For simplicity, no collision detection for the objects was enforced or deemed necessary. The graphics calls are communicated from the behavior engine on the PC over Ethernet to a Performer-based graphics system running on an 8-processor Silicon Graphics Onyx2 Infinite Reality system. A constant rendering rate of 30 FPS was achieved. The graphics were output to a NEC LCD projector displaying to a 5x4 screen (See Figure 44) positioned seven feet in front of the user.

### **6.2.3 Perception and action**

The categorical framework for movement recognition described in this research has a direct impact in general to artificial life domains. The categorical models and their relation to natural movements and communication in nature can benefit the artificial life community by offering a well-motivated computational model for movement perception and behavior of artificial creatures. The motion models developed here can also be applied to enable life-like movements of creatures within artificial worlds. Since the underlying model for the motion category is a generative model, it could also be used to synthesize movements for the artificial creatures. Thus we present the idea of a *shared* model for performing and recognizing motions. Here the connection between “doing” and “seeing” is direct, relating at the representation level the possible generation of creature motions to those patterns that are perceivable. A future demonstration related to the virtual worlds could entail the artificial creatures embodying a single model connected to their motor performance and their perceptual recognition. Thus the person could be removed from the current hummingbird world and the generative model could be used to autonomously control the hummingbird motions and perceptions. This approach would advance the system towards a more autonomous environment that can be explored systematically with action-perception behaviors.

## **6.3 Section summary**

In this section, we integrated the categorical recognition models within an interactive virtual world populated by artificial hummingbirds. The system used the infrared tracking apparatus to give control of a male hummingbird to the user. The recognition system was employed on the point-light tracking data and the results were connected to the behavior system of a watching female hummingbird. Then, as the user performed certain motion patterns, the recognition system

directed the motivation systems for the hummingbirds to react with special behaviors. This visual experience (highlighted in [57]) was used to show that the category models can be incorporated within interactive virtual worlds.

## 7 Conclusion and Summary

In this thesis, we presented a categorical computational model and implementation for an oscillatory motion perception system. The significance of the categorical approach was first illustrated in the appearance of categorical oscillatory motions used throughout the animal kingdom as a form of communication (Claim 1), and by a single sinusoidal model (Claim 2) that can be used to hierarchically organize the motions to specify the types of computation required for recognition (Claim 3). More specifically, the approach promotes collecting an important set of related motions, organizing a hierarchical complexity ordering based on the number of parameters and constraints of a characterizing model, deriving sub-categorizations showing the prominent specializations, and using the resulting sub-categorical organization as the specification for local computation in the recognition implementation. This method was demonstrated with examples of synthetic and real motions, removing the reliance on constructing individual ad hoc models for recognition. This computational approach further exemplifies the notion that there is structure and regularity in the world (motions) that in turn make categorical representations and perceptual mechanisms reliable at interpreting meaningful patterns in nature.

We gain insight into the motions deemed important by nature by observing that certain types of patterns appear across species. Such categorical motion patterns reflect regularities in nature and provide useful indicators about animal intentions, perhaps even revealing their emotional state (more generally addressed by [29]). In this research we have shown that oscillatory motions in particular are quite prevalent throughout the animal kingdom as used for communication. These oscillatory motions include *up-and-down*, *side-to-side*, *circle*, *spiral*, *loop*, *undulate*, *figure-8*, and *U-shuttle*. All of these motions can be characterized using a simple sinusoidal model with very specific and limited parameter values. For these motions, the main distinctive structural regularities for the parameters are frequency ratios of  $\{1:1, 1:2\}$  and relative phases of  $\{0, \pm\pi/2, \pi\}$ . An allowance for separate stylistic variation in the patterns was also presented.

We showed that these motions could be ordered in complexity using the notion of increasing specialization in the model parameters, where simpler motions are defined as those patterns requiring fewer parameters or constraints. The result was a hierarchical representation of the motions with the simplest motions at the top and the most complex motions at the bottom. A sub-categorical partitioning of this hierarchy revealed the prominent forms of oscillatory

specialization in the motions. The sub-categories listed from simple to complex are 1-D motion (up-and-down, side-to-side), 2-D motion (circle), translation (spiral, loop, undulate), and frequency-doubling (U-shuttle, figure-8). To show a basic correlation of this ordering to observations of biological motion, we presented a variety of species performing the simplest motions under the assertion that the simplest movements should be the most ubiquitous across species. Furthermore, a large sampling of different bird species suggests that while particular species prefer performing certain characteristic patterns, the simpler 1-D and 2-D movements tend to have a wider usage across birds. Additionally, the partial ordering for the circle, U-shuttle, and figure-8 motions as derived by the model was identified in the communicative dance language of honey bees.

The categorical structure for the motions was then exploited by a perceptual model designed to recognize these patterns. The structural regularities within levels (sub-categories) of complexity in the hierarchy for the motions were implemented as a series of locally organized and ranked computations. Here, each sub-category has localized computation to recognize the components unique to its group. The structure of the category is therefore reflected in the evolution of computation for recognition. The recognition approach was composed of three main stages. First, single pattern cycles were extracted from the input position trajectories of the visually-tracked object. The position trajectories of arm motion were determined by experiment to be more stable and desirable than the corresponding velocity trajectories to represent the pattern. Sign-changes (zero-crossings) in the appropriate velocity trajectory were used to segment single pattern cycles, with recognition taking place at these zero-crossing locations. Second, Fourier analysis was used to extract the appropriate sinusoidal model parameters (of amplitude, frequency, phase) from the position trajectories. Third, the estimated motion parameters were categorically compared with the models looking for a structural match (and possibly a stylistic comparison) using either full category recognition (checking all models in the category) or goal-directed recognition (examining only specific models).

The approach was demonstrated with experiments using synthetic, point-light, and video motions. To examine the nature of confusions caused from the differing complexities of the models, synthetic and user-generated point-light motions were tested against each model. We found that slight confusions may occur for some related models (e.g. circle, loop) from the existence (or not) of a more complex component in the signal. The 3-D spiraling motion was also shown to be more difficult to recognize. Results had a few missed pattern cycles (or portions of a pattern cycle). The

errors in these cases were mainly due to phase and translation estimations beyond accepted ranges. Lastly, oscillations from real video scenes were tested to show the applicability of the approach with non-engineered input. Video sequences of a person nodding and waving his head, a classic animation sequence with an object looping through the air, and a movie of a gibbon swinging across a rope (using a newly derived swinging model) were used to show the strength of the recognition system.

To further explore the category models, we presented an interactive virtual world where the recognition of the oscillatory motions affected the behaviors of artificial hummingbirds. Motivated by interactive and autonomous virtual environments, the virtual hummingbird world enabled a user to control the motions of a male hummingbird trying to convey its intentions to a watching female. The user had direct control over the motions of the male hummingbird by using the infrared tracking system. If the user repeatedly performed one of four categorical oscillations (up-and-down, side-to-side, figure-8, and U-shuttle), the female hummingbird responded appropriately with behaviors of acceptance, fear, and courtship. This demonstration illustrated the applicability of the recognition system within an interactive entertaining experience.

## ***7.2 Significance and contributions***

The general significance of this research, apart from the inherent intellectual interest, is widespread. For the vision and computer scientist, this research yields a strong motivation for the unification of computational methods used for recognizing related motion patterns. This work provides a computational strategy for recognizing such movements based on categories, rather than employing a series of ad hoc or training-based methods for recognition. For the cognitive scientist, the results of this work show a simple modeling and structuring which underlies specialized cognitive patterns. The work derived from this study also offers insight for the ethologist into a structured commonality of movement and behavior patterns across animals. Given the fundamental perceptual, cognitive, and behavioral nature of this work, any insight into the form, function, and recognition of motion patterns will add significance to the understanding of perception in general and to the design of machines that are to interact in the world with humans, animals, or other artificial creatures.

The main contribution in terms of computational significance is that we developed a method to specify the types and layout of computation for recognition based on the underlying categorical motion organization. A common, and alternative, recognition approach is to use training-based methods to build prototypical models where a full computational load is assigned and applied to each input motion. As discussed in Section 1.2, these methods ignore the categorical foundation underlying structurally related motions. Instead we developed an approach to specify and organize computation using a hierarchical ranking of the motions based on the limited parameter values and constraints within the category to organize the motions. This ranking can be used to partition the motions into sub-categorical specializations (separable components of the model that are added to generate complex patterns from simpler patterns). The hierarchical sub-categories suggest a more general layout for recognition, where computation is locally assigned to the specializations within the class of motion. The result is a more general recognition framework with simpler motions needing only small amounts of computation and more complex motions requiring additional computation. A further contribution in terms of this organization is that we studied several species of bird and found that the oscillatory motions we deemed as simple from the model ranking were found to be generally more ubiquitous across birds. A representative sampling of bird species and their oscillatory motions are made available in the appendices of this thesis. We have not found such a focussed compilation of pattern-related display motions in the literature. Additionally, the specific ordering for three of the motions (circle, figure-8, U-shuttle) was identified in the transformations of the dance language in honey bees.

Another contribution of this research was presented in the experimental results that show human arm oscillations are better represented with position data rather than higher-order derivatives. Previous experiments and research promote the use of “minimum-jerk” models to represent simple arm/hand movements. These methods minimize the third derivative of position (“jerk”) using average data to obtain a representative (and predictive) form of the motion pattern. These experiments use simple reaching, pointing, and turning arm movements. We experimentally showed under various performance conditions that human arm performance of these more *cognitively-based* oscillatory patterns shows considerable derivative variance whereas the position information remains quite stable. Thus the contribution from this experiment is that the oscillatory patterns appear to not be derivative-minimized as are the other goal-directed arm movements.

Lastly, a real-time computer implementation was outlined and demonstrated that recognizes the oscillatory patterns based on a structured categorical organization for the motions, computation, and recognition. The implementation served as a means to demonstrate the model, testing the use of a sub-categorical assignment of computation for reliable recognition with real oscillatory motion. This approach shows the applicability of a categorical assignment of computation without the need for statistical training on large sets of examples.

### **7.3 Further questions and problems with the approach**

Due to the large breadth of research for this thesis, further questions and problems with the approach will naturally arise. Multiple issues exist with respect to speculations pertaining to the generative-perceptual model, pre-knowledge for category formulation, data representation, and biological similarity.

First, there is a concern that we don't really know for this set of oscillatory motions if animals are truly responding to the actual motion patterns. For example, a *color-flash* during movement or perhaps a repetitive *approach* may be the true perceptual features rather than the complete motion pattern. From a generative standpoint, the motions are most likely represented as complete patterns, but this not necessarily true for their perception. Our approach couples the action and perception into a single model. We gathered our set of oscillatory motions mainly from verbal descriptions in the literature describing animal behaviors. Thus at least we, as observers of these motions, identify these particular and stereotypical patterns. Perhaps this research partially addresses how humans perceive and recognize motion patterns.

Our approach also requires an initial compilation of a set of related motions to define and organize the category. Here one can question whether prior knowledge of the category is required to build the category. This is not problematic if we wish to define a computational structure for an existing set of motions. Additionally, given the basic sub-categories (for the most primitive motions), the category could be *evolved* to accommodate the later more complex patterns. If we consider this case, then an explanation is required of what drives the extensions to the category with particular parameters and constraints (e.g. from biological limitations or mechanisms, reliability of communication).

In terms of the motion data representation, we analyzed motions that are tracked based on a single feature point (e.g. the body in motion, a color patch on a moving body part). This point-tracking approach made it simple to focus the effort on categorizing the oscillatory motion trajectories. How well the categorical approach extends to other non-point representational forms is not easy to answer. For example, in Section 2.2 we briefly mentioned a previous approach to recognition based on a “template” of motion that gathers in a single image the motions throughout an action. This image-based approach to representing the motion pattern is quite different than using trajectories. It is not clear how such image representations can be similarly characterized and organized as the sinusoidal trajectory motions. Perhaps these non-trajectory representations can segment the motion patterns into regions or “patches” of coherent motion which can then be either tracked (resulting in trajectory-based categorization) or categorized based on spatial (e.g. symmetry) and temporal (e.g. before, during, after) relations.

We would also like a stronger correlation of the model-based motion ordering to biological or psychological orderings to further support the approach. Our current comparison with several bird species is suggestive of simple and complex sub-categories, but it is not strong enough to validate the overall hierarchical motion organization. We have shown a plausible sub-ordering for circle, figure-8, and U-shuttle with the transitional lattice for the honey bee dances, and we would prefer this type of result (mapping) for the entire class of motions. Perhaps a stronger biological correspondence for the motions would help to resolve the sub-categorization for undulate motion (belonging to the translation sub-category, but actually residing at the previous hierarchical level). Exploring a psychophysical correlation would be another approach to validate or suggest a model ordering, experimenting with human perceptual categorization to programmed motions.

## **7.4 *Future directions***

Future work involving categorical motion perception includes looking deeper at the stylistic parameters, looking for additional categories of motion, and comparing the categories to cognitive and biological systems. In this research, we have explored the structural components for the class of oscillatory motions. It is of future interest to examine how different species (including humans) use the stylistic component to alter the meaning or expressiveness of the patterns. For example, nodding the head vigorously has a stronger emphasis than a casual nod. Also, the topic of this research was focussed on the specific category of oscillatory motions. Preliminary research



reveals that additional categories of communicative motion exist across animals. For example, as presented earlier, motions best described as *impulse* and *stochastic/vibratory* can be readily found across species (See examples in Appendix C). It is reasonable to expect structural regularities in these categories as well (e.g. directions and extents of impulse, speed of vibration) to help define categorical perception models for their recognition. As shown with our comparisons of the complexity ranking for the oscillations to observations of animal movement, additional comparisons of derived categorical structures with psychophysical experiments (where is the visual boundary between swaying and vibration?) and biological systems (evolutionary correspondences?) would further justify the use of categories as a means for designing perceptual, and perhaps motor, mechanisms.

The manner of recognition for more complex and composite (multiple) motions is also of interest for future research. Such motions are comprised of multiple motion *pieces*, and therefore require a more involved recognition approach compared to the single point-light trajectories. In this case, the individual motions may be from the same class or utterly different categories. For example, the ostrich performs its courtship display by alternately lifting up-and-down its wings while tossing his head side-to-side [10]. In this example, we see three individual oscillatory motions performed. Another example is with the Grey-headed kingfisher, where its display consists of pivoting side-to-side while vibrating its wings [44]. Here there is an oscillatory side-to-side motion with the addition of a “fluttering” vibration. These two motions most likely belong to two different motion categories. A last example is of a Blue grouse who for its display performs a flutter-jump (See Appendix C). In this motion, the bird leaps into the air while wildly shaking and fluttering its entire body [32]. Again we see the shaking motion but this time accompanied by an impulse-like jumping action. So clearly there are motions that not only consist of patterns within the same category, but also are comprised of multiple motion categories – a categorical cross-over. Such composite motions are not unique to animal or biological motion, but instead can be related to other moving objects in our world. For example, cars and bicycles have an overall translation component with “spinning” wheels. We could even attempt to represent a person dribbling a basketball in the same fashion (motion of person with the up-and-down motion of the ball). This may perhaps extend to the interaction between multiple people, as in a coordinated dance routine where the *overall* appearance pattern is produced by the combination of individual motions of the couple.

Though we did not address this manner of composite recognition in this thesis, we speculate that a higher-level piece-wise recognition process using the categories is likely. For instance, a larger composite motion could first be decomposed into its component categorical pieces, where each individual motion is then recognized categorically. The higher-level representation for the composite motion may have a stored topological layout signifying spatial relations between the various component motions. For example with the ostrich, this layout may specify *up-and-down motions* on the left and right side, and a *side-to-side motion* in the middle-top region. This layout is triangular with three components all belonging to an oscillation category. With a body-centered coordinate frame, this would enable location invariance (placement in the image) during recognition. For the more “blended” motions as in a flutter-jump, the impulse jump would most likely have to be de-coupled and separated from the stochastic fluttering motion. This may be attained by sequential recognition of the component motions. We could first identify the jump, then remove that impulse component leaving the fluttering motion left to be recognized. This may be an approach for the characterization of the moving car motion as well – first the translation of the car is removed and then the spinning/turning wheels are identified (with a spatial layout in reference to the car). We offer these basic descriptive approaches to recognition of more complex composite motions as a future research emphasis.

This categorical approach was demonstrated with a class of oscillatory motions prominently used in communicative bird behavior. The resultant framework (organized structural regularities) is not confined to only this class of motion, but rather can be generalized to include other classes of motion. We find that these same types of motion patterns used in communication also appear in other non-communicative behaviors. Many of these communicative behaviors of animals are derived (ritualized) from more basic actions. The communicative motions and recognition systems are likely to be co-evolved from the more basic behavioral mechanisms, with specialization of the general perceptual system to address the derived communicative acts. Therefore our categorical approach characterizing the structure in communicative patterns gives some insight into the more general motion perception system for animals that require the identification of movements throughout their daily activities. Several goals of an animal’s motion perception system (including interaction with prey, predators, and conspecifics) are thus addressed with this categorical framework in that these models can be employed for the identification and recognition of both communicative and non-communicative behaviors. Additionally, if we consider the recognition of composite motions having categorical components (as described above), the models are applicable to the interpretation of an even broader range of

motions (including in-animate motions). Thus our categorical approach to the characterization of communicative motion patterns is most parsimonious in that it contributes to the larger tasks required of a more general motion perception machine.

As a result of our research, we find that several important communicative motions in nature can be categorically represented and recognized. The approach uses a simple model within a hierarchical organization to identify the structural foundation. That this categorical structure can be used to design artificial perception systems to recognize these motions was shown using synthetic and real motions. The approach was additionally demonstrated with an interactive virtual world, where the categorical recognition models were used to affect the behaviors of virtual creatures. Such a categorical description and organization for motion is useful for understanding the generation and recognition of structurally related movement patterns. Understanding how these categories are structured may give insight into how we, ourselves, categorize and develop our own cognitive models for patterns of motion in nature.

## **8 Acknowledgements**

I would first like to acknowledge my research supervisors Aaron Bobick and Whitman Richards. Their comments, suggestions, and creative influence throughout this research impacted the way I now think about perception (both human and machine). I would also like to thank my thesis committee members, Bruce Blumberg, Jim Gould, and Demetri Terzopoulos, for their helpful comments and suggestions for this thesis. Having such a multi-disciplinary thesis and committee was indeed an interesting and engaging experience. For their help in putting together the virtual hummingbird world, I thank Bill Tomlinson and Jed Wahl. Lastly, I would like to thank Tiffany Thornton and my family for their support and encouragement throughout the years.

## Appendix A

This appendix lists the birds used for determining the *across*-species preference of the oscillatory motions. We initially examined several hundred bird species to determine this particular set of oscillatory motions. Here we present a smaller, yet representative, sampling of those species. Fourteen types of birds are listed in alphabetical order. For each motion listed in the species tables, a verbal description, species name (English and scientific), and reference are given.

| <b>Ducks</b>          |                              |   |                  |
|-----------------------|------------------------------|---|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b>    | <b>Species</b>                              | <b>Reference</b> |
| Up-and-down           | Head pumping                 | Mallard<br>( <i>Anas platyphynchos</i> )    | [91]             |
| Side-to-side          | Bill moved from side to side | Wigeon<br>( <i>Anas penelope</i> )          | [91]             |
| Circle                | Circle female                | Smew<br>( <i>Mergus albellus</i> )          | [91]             |
| Spiral                | [Unavailable]                | [Unavailable]                               | [Unavailable]    |
| Loop                  | [Unavailable]                | [Unavailable]                               | [Unavailable]    |
| Undulate              | [Unavailable]                | [Unavailable]                               | [Unavailable]    |
| Figure-8              | Waving bill in figure-8      | Ruddy duck<br>( <i>Oxyura jamaicensis</i> ) | [91]             |
| U-shuttle             | [Unavailable]                | [Unavailable]                               | [Unavailable]    |

| <b>Falcons</b>        |                             |   |                  |
|-----------------------|-----------------------------|---|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b>   | <b>Species</b>                                  | <b>Reference</b> |
| Up-and-down           | Bob head rapidly            | Kestrel<br>( <i>Falco tinnunculus</i> )         | [92]             |
| Side-to-side          | Side to side swing          | Peregrine falcon<br>( <i>Falco peregrinus</i> ) | [92]             |
| Circle                | Both circling               | Kestrel<br>( <i>Falco tinnunculus</i> )         | [92]             |
| Spiral                | High-circling far up in sky | Peregrine falcon<br>( <i>Falco peregrinus</i> ) | [92]             |
| Loop                  | Loop-the-loops              | Peregrine falcon<br>( <i>Falco peregrinus</i> ) | [96]             |
| Undulate              | Undulating flight           | Peregrine falcon<br>( <i>Falco peregrinus</i> ) | [96]             |
| Figure-8              | Figure-8 flight             | Peregrine falcon<br>( <i>Falco peregrinus</i> ) | [92]             |
| U-shuttle             | Swings low above in arcs    | Hobby<br>( <i>Falco subbuteo</i> )              | [92]             |

## Gulls

| Motion Pattern | Motion Description     | Species  | Reference     |
|----------------|------------------------|--|---------------|
| Up-and-down    | Head pumping           | Black-headed gull<br>( <i>Larus ridibundus</i> ) | [90]          |
| Side-to-side   | Turn from side to side | Black-headed gull<br>( <i>Larus ridibundus</i> ) | [90]          |
| Circle         | Circular flight        | Black-headed gull<br>( <i>Larus ridibundus</i> ) | [90]          |
| Spiral         | [Unavailable]          | [Unavailable]                                    | [Unavailable] |
| Loop           | [Unavailable]          | [Unavailable]                                    | [Unavailable] |
| Undulate       | [Unavailable]          | [Unavailable]                                    | [Unavailable] |
| Figure-8       | [Unavailable]          | [Unavailable]                                    | [Unavailable] |
| U-shuttle      | [Unavailable]          | [Unavailable]                                    | [Unavailable] |

## Hawks

| Motion Pattern | Motion Description       | Species   | Reference     |
|----------------|--------------------------|---|---------------|
| Up-and-down    | Bobs head up and down    | Broad-winged hawk<br>( <i>Buteo platypterus</i> ) | [96]          |
| Side-to-side   | [Unavailable]            | [Unavailable]                                     | [Unavailable] |
| Circle         | High circling            | Sparrowhawk<br>( <i>accipiter nisus</i> )         | [92]          |
| Spiral         | Soaring in tight spirals | Sparrowhawk<br>( <i>accipiter nisus</i> )         | [92]          |
| Loop           | [Unavailable]            | [Unavailable]                                     | [Unavailable] |
| Undulate       | Undulating flight        | Sparrowhawk<br>( <i>accipiter nisus</i> )         | [92]          |
| Figure-8       | [Unavailable]            | [Unavailable]                                     | [Unavailable] |
| U-shuttle      | [Unavailable]            | [Unavailable]                                     | [Unavailable] |

| <b>Hérons</b>         |                           |   |                  |
|-----------------------|---------------------------|---|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b> | <b>Species</b>                                | <b>Reference</b> |
| Up-and-down           | Pumps bill up and down    | Squacco heron<br>( <i>Ardeola ralloides</i> ) | [91]             |
| Side-to-side          | Swaying from side to side | Purple heron<br>( <i>Ardea purpurea</i> )     | [91]             |
| Circle                | Brief circle flights      | Grey heron<br>( <i>Ardea cinerea</i> )        | [91]             |
| Spiral                | [Unavailable]             | [Unavailable]                                 | [Unavailable]    |
| Loop                  | [Unavailable]             | [Unavailable]                                 | [Unavailable]    |
| Undulate              | [Unavailable]             | [Unavailable]                                 | [Unavailable]    |
| Figure-8              | [Unavailable]             | [Unavailable]                                 | [Unavailable]    |
| U-shuttle             | [Unavailable]             | [Unavailable]                                 | [Unavailable]    |

| <b>Hummingbirds</b>   |   |  |                  |
|-----------------------|---|--|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b>                 | <b>Species</b>   | <b>Reference</b> |
| Up-and-down           | Ascend and descend                        | Ruby-throated hummingbird<br>( <i>Archilochus colubris</i> )   | [96]             |
| Side-to-side          | Back and forth horizontally               | Ruby-throated hummingbird<br>( <i>Archilochus colubris</i> )   | [96]             |
| Circle                | Circles territory                         | Broad-tailed hummingbird<br>( <i>Selasphorus platycercus</i> ) | [97]             |
| Spiral                | Vertical spiraling flight                 | Lucifer hummingbird<br>( <i>Calothorax lucifer</i> )           | [32]             |
| Loop                  | Alternating looping                       | White-eared hummingbird<br>( <i>Hylocharis leucotis</i> )      | [32]             |
| Undulate              | [Unavailable]                             | [Unavailable]  | [Unavailable]    |
| Figure-8              | Narrow figure-8 pattern                   | Black-chinned hummingbird<br>( <i>Archilochus alexandri</i> )  | [32]             |
| U-shuttle             | Flies precise path of a wide U-shaped arc | Ruby-throated hummingbird<br>( <i>Archilochus colubris</i> )   | [96]             |

| <b>Kingfishers</b>    |                            |  |                  |
|-----------------------|----------------------------|--|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b>  | <b>Species</b>   | <b>Reference</b> |
| Up-and-down           | Tail bobbing               | Kingfisher<br>( <i>Alcedo atthis</i> )                     | [17]             |
| Side-to-side          | Sways from side to side    | Kingfisher<br>( <i>Alcedo atthis</i> )                     | [17]             |
| Circle                | Flying in circles          | Kingfisher<br>( <i>Alcedo atthis</i> )                     | [17]             |
| Spiral                | Descending in tight spiral | White-breasted kingfisher<br>( <i>Halcyon smyrnensis</i> ) | [17]             |
| Loop                  | [Unavailable]              | [Unavailable]  | [Unavailable]    |
| Undulate              | [Unavailable]              | [Unavailable]  | [Unavailable]    |
| Figure-8              | [Unavailable]              | [Unavailable]  | [Unavailable]    |
| U-shuttle             | [Unavailable]              | [Unavailable]  | [Unavailable]    |

| <b>Larks</b>          |                                   |  |                  |
|-----------------------|-----------------------------------|--|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b>         | <b>Species</b>                               | <b>Reference</b> |
| Up-and-down           | Bobbing movements                 | Woodlark<br>( <i>Lullula arborea</i> )       | [16]             |
| Side-to-side          | Flies horizontally back and forth | Thekla lark<br>( <i>Galerida Theklae</i> )   | [16]             |
| Circle                | Circles at fairly constant height | Woodlark<br>( <i>Lullula arborea</i> )       | [16]             |
| Spiral                | Spirals up                        | Woodlark<br>( <i>Lullula arborea</i> )       | [16]             |
| Loop                  | Irregular loops                   | Woodlark<br>( <i>Lullula arborea</i> )       | [16]             |
| Undulate              | Undulations                       | Crested lark<br>( <i>Galerida cristata</i> ) | [16]             |
| Figure-8              | [Unavailable]                     | [Unavailable]                                | [Unavailable]    |
| U-shuttle             | [Unavailable]                     | [Unavailable]                                | [Unavailable]    |



| <b>Owls</b>           |                           |   |                  |
|-----------------------|---------------------------|---|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b> | <b>Species</b>                              | <b>Reference</b> |
| Up-and-down           | Bobs body up and down     | Little owl<br>( <i>Athene noctua</i> )      | [17]             |
| Side-to-side          | Sway head side to side    | Snowy owl<br>( <i>Nyctea scandiaca</i> )    | [17]             |
| Circle                | Circling nest site        | Long-eared owl<br>( <i>Asio otus</i> )      | [17]             |
| Spiral                | Spirals upwards           | Great grey owl<br>( <i>Strix nebulosa</i> ) | [17]             |
| Loop                  | [Unavailable]             | [Unavailable]                               | [Unavailable]    |
| Undulate              | Undulating flight         | Snowy owl<br>( <i>Nyctea scandiaca</i> )    | [33]             |
| Figure-8              | [Unavailable]             | [Unavailable]                               | [Unavailable]    |
| U-shuttle             | [Unavailable]             | [Unavailable]                               | [Unavailable]    |

| <b>Pigeons</b>        |                           |   |                  |
|-----------------------|---------------------------|---|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b> | <b>Species</b>                                    | <b>Reference</b> |
| Up-and-down           | Rapid bobbing             | Turtle dove<br>( <i>Streptopelia turtur</i> )     | [17]             |
| Side-to-side          | [Unavailable]             | [Unavailable]                                     | [Unavailable]    |
| Circle                | Small circular flight     | Namaqua dove<br>( <i>Oena capensis</i> )          | [17]             |
| Spiral                | Glides in spiral          | Collared dove<br>( <i>Streptopelia decaocto</i> ) | [17]             |
| Loop                  | [Unavailable]             | [Unavailable]                                     | [Unavailable]    |
| Undulate              | [Unavailable]             | [Unavailable]                                     | [Unavailable]    |
| Figure-8              | [Unavailable]             | [Unavailable]                                     | [Unavailable]    |
| U-shuttle             | [Unavailable]             | [Unavailable]                                     | [Unavailable]    |

| <b>Plovers</b>        |                           |  |                  |
|-----------------------|---------------------------|--|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b> | <b>Species</b>                                       | <b>Reference</b> |
| Up-and-down           | Alarm bobbing             | Kentish plover<br>( <i>Charadrius alexandrinus</i> ) | [90]             |
| Side-to-side          | Tilting from side to side | Kentish plover<br>( <i>Charadrius alexandrinus</i> ) | [90]             |
| Circle                | Circling                  | Grey plover<br>( <i>Pluvialis squatarola</i> )       | [90]             |
| Spiral                | Descends in circles       | Caspian plover<br>( <i>Charadrius asiaticus</i> )    | [90]             |
| Loop                  | [Unavailable]             | [Unavailable]  | [Unavailable]    |
| Undulate              | [Unavailable]             | [Unavailable]  | [Unavailable]    |
| Figure-8              | Flies in figure-8         | Grey plover<br>( <i>Pluvialis squatarola</i> )       | [90]             |
| U-shuttle             | [Unavailable]             | [Unavailable]  | [Unavailable]    |

| <b>Terns</b>          |                           |  |                  |
|-----------------------|---------------------------|--|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b> | <b>Species</b>                         | <b>Reference</b> |
| Up-and-down           | Head bobbing              | Swift tern<br>( <i>Sterna bergii</i> ) | [17]             |
| Side-to-side          | Sways from side to side   | Swift tern<br>( <i>Sterna bergii</i> ) | [17]             |
| Circle                | Strut in tight circle     | Swift tern<br>( <i>Sterna bergii</i> ) | [17]             |
| Spiral                | Spirals upward            | Swift tern<br>( <i>Sterna bergii</i> ) | [17]             |
| Loop                  | [Unavailable]             | [Unavailable]                          | [Unavailable]    |
| Undulate              | [Unavailable]             | [Unavailable]                          | [Unavailable]    |
| Figure-8              | [Unavailable]             | [Unavailable]                          | [Unavailable]    |
| U-shuttle             | [Unavailable]             | [Unavailable]                          | [Unavailable]    |

| <b>Wheatears</b>      |                             |  |                  |
|-----------------------|-----------------------------|--|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b>   | <b>Species</b>                                     | <b>Reference</b> |
| Up-and-down           | Wagging up and down of tail | Desert wheatear ( <i>Oenanthe deserti</i> )        | [16]             |
| Side-to-side          | Head swaying                | Black-eared wheatear ( <i>Oenanthe hispanica</i> ) | [16]             |
| Circle                | Circles territory           | Desert wheatear ( <i>Oenanthe deserti</i> )        | [16]             |
| Spiral                | Spirals downward            | Isabelline wheatear ( <i>Oenanthe isabellina</i> ) | [33]             |
| Loop                  | [diagram]                   | Desert wheatear ( <i>Oenanthe deserti</i> )        | [16]             |
| Undulate              | [Unavailable]               | [Unavailable]                                      | [Unavailable]    |
| Figure-8              | [diagram]                   | Desert wheatear ( <i>Oenanthe deserti</i> )        | [16]             |
| U-shuttle             | [Unavailable]               | [Unavailable]                                      | [Unavailable]    |

| <b>Woodpeckers</b>    |                              |   |                  |
|-----------------------|------------------------------|---|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b>    | <b>Species</b>  | <b>Reference</b> |
| Up-and-down           | Head bobbing                 | Pileated woodpecker ( <i>Dryocopus pileatus</i> )       | [32]             |
| Side-to-side          | Head swaying                 | Grey-headed woodpecker ( <i>Picus canus</i> )           | [17]             |
| Circle                | Flying in wide circles       | Syrian woodpecker ( <i>Dendrocopos syriacus</i> )       | [17]             |
| Spiral                | Climbing up spiraling        | Middle spotted woodpecker ( <i>Dendrocopos medius</i> ) | [17]             |
| Loop                  | [Unavailable]                | [Unavailable]   | [Unavailable]    |
| Undulate              | Undulating flight            | Yellow-bellied sapsucker ( <i>Sphyrapicus varius</i> )  | [32]             |
| Figure-8              | Bill describes figure-8 path | Syrian woodpecker ( <i>Dendrocopos syriacus</i> )       | [17]             |
| U-shuttle             | [Unavailable]                | [Unavailable]   | [Unavailable]    |

## Appendix B

This appendix lists the birds used for determining the *individual*-species preference of the oscillatory motions. Listed are various species of hummingbird, falcon, and woodpecker. For each motion listed in the species tables, a verbal description, species name (English and scientific), and reference are given.

| <b>Falcons</b>        |                           |   |                  |
|-----------------------|---------------------------|---|------------------|
| <b>Motion Pattern</b> | <b>Motion Description</b> | <b>Species</b>                                    | <b>Reference</b> |
| Up-and-down           | Bob head                  | Kestral<br>( <i>Falco tinnunculus</i> )           | [92]             |
|                       | Rapid head bobbing        | Eleonora's falcon<br>( <i>Falco eleonora</i> )    | [92]             |
|                       | Head/body bobbing         | Peregrine falcon<br>( <i>Falco peregrinus</i> )   | [92]             |
| Side-to-side          | Side to side swing        | Peregrine falcon<br>( <i>Falco peregrinus</i> )   | [92]             |
| Circle                | Circle flight             | Lesser kestrel<br>( <i>Falco naumanni</i> )       | [92]             |
|                       | Circling                  | Kestrel<br>( <i>Falco tinnunculus</i> )           | [92]             |
|                       | Circling                  | Red-footed falcon<br>( <i>Falco vespertinus</i> ) | [92]             |
|                       | Circle                    | Merlin<br>( <i>Falco columbarius</i> )            | [92]             |
|                       | Circle flights            | Hobby<br>( <i>Falco subbuteo</i> )                | [92]             |
|                       | High circling             | Lanner<br>( <i>Falco biarmicus</i> )              | [92]             |
|                       | High circling             | Saker<br>( <i>Falco cherrug</i> )                 | [92]             |
|                       | High circling             | Gyr Falcon<br>( <i>Falco rusticolus</i> )         | [92]             |
|                       | High circling             | Peregrine falcon<br>( <i>Falco peregrinus</i> )   | [33]             |
| Spiral                | Spirals                   | Peregrine falcon<br>( <i>Falco peregrinus</i> )   | [92]             |
| Loop                  | Looping the loop          | Peregrine falcon<br>( <i>Falco peregrinus</i> )   | [92]             |
| Undulate              | Wave-like motion          | Red-footed falcon<br>( <i>Falco vespertinus</i> ) | [92]             |
|                       | Undulation flight         | Peregrine falcon<br>( <i>Falco peregrinus</i> )   | [96]             |

|           |                    |   |      |
|-----------|--------------------|---|------|
| Figure-8  | Figure-8           | Hobby<br>( <i>Falco subbuteo</i> )              | [92] |
|           | Figure-8           | Peregrine falcon<br>( <i>Falco peregrinus</i> ) | [92] |
| U-shuttle | Swings low in arcs | Hobby<br>( <i>Falco subbuteo</i> )              | [92] |

| Hummingbirds   |                             |  |               |
|----------------|-----------------------------|--|---------------|
| Motion Pattern | Motion Description          | Species  | Reference     |
| Up-and-down    | Ascend and descend          | Ruby-throated hummingbird<br>( <i>Archilochus colubris</i> )   | [96]          |
| Side-to-side   | Back and forth horizontally | Ruby-throated hummingbird<br>( <i>Archilochus colubris</i> )   | [96]          |
|                | Rotates from side to side   | Anna's hummingbird<br>( <i>Calypte anna</i> )                  | [97]          |
| Circle         | Circles territory           | Broad-tailed hummingbird<br>( <i>Selasphorus platycercus</i> ) | [97]          |
|                | Flying in high circle       | Anna's hummingbird<br>( <i>Calypte anna</i> )                  | [97]          |
| Spiral         | Vertical spiraling flight   | Lucifer hummingbird<br>( <i>Calothorax lucifer</i> )           | [32]          |
|                | Slow spiraling              | Allen's hummingbird<br>( <i>Selasphorus rufus</i> )            | [97]          |
| Loop           | Alternating looping         | White-eared hummingbird<br>( <i>Hylocharis leucotis</i> )      | [32]          |
| Undulate       | [Unavailable]               | [Unavailable]  | [Unavailable] |
| Figure-8       | Narrow figure-8 pattern     | Black-chinned hummingbird<br>( <i>Archilochus alexandri</i> )  | [32]          |

|           |   |  |      |
|-----------|---|--|------|
| U-shuttle | Flies precise path of a wide U-shaped arc | Ruby-throated hummingbird<br>( <i>Archilochus colubris</i> )   | [96] |
|           | Pendulum-like swings                      | Lucifer hummingbird<br>( <i>Calothorax lucifer</i> )           | [32] |
|           | Pendulum-swing                            | Broad-billed hummingbird<br>( <i>Cynanthus latirostris</i> )   | [32] |
|           | Swings pendulumlike                       | Black-chinned hummingbird<br>( <i>Archilochus alexandri</i> )  | [32] |
|           | Flies in U-shaped pattern                 | Costa's hummingbird<br>( <i>Calypte costae</i> )               | [32] |
|           | Flies in short tight arcs                 | Anna's hummingbird<br>( <i>Calypte anna</i> )                  | [97] |
|           | Flies in U-shaped pattern                 | Calliope hummingbird<br>( <i>Stellula calliope</i> )           | [32] |
|           | Flies in U-shaped pattern                 | Broad-tailed hummingbird<br>( <i>Selasphorus platycercus</i> ) | [97] |
|           | Pendulumlike rocking                      | Allen's hummingbird<br>( <i>Selasphorus rufus</i> )            | [97] |

| Woodpeckers    |                     |   |           |
|----------------|---------------------|---|-----------|
| Motion Pattern | Motion Description  | Species   | Reference |
| Up-and-down    | Head bobbing        | Pileated woodpecker<br>( <i>Dryocopus pileatus</i> )        | [32]      |
|                | Head bobbing        | Northern flicker<br>( <i>Colaptes auratus</i> )             | [32]      |
|                | Head bobbing        | Gila woodpecker<br>( <i>Melanerpes uropygialis</i> )        | [32]      |
|                | Head bobbing        | Williamson's sapsucker<br>( <i>Sphyrapicus thyroideus</i> ) | [32]      |
|                | Head bobbing        | Ladder-backed woodpecker<br>( <i>Picoides scalaris</i> )    | [32]      |
|                | Head bobbing        | Red-cockaded woodpecker<br>( <i>Picoides borealis</i> )     | [32]      |
|                | Head bobbing        | Nuttall's woodpecker<br>( <i>Picoides nuttallii</i> )       | [32]      |
|                | Bobbing             | Grey-headed woodpecker<br>( <i>Picus canus</i> )            | [17]      |
|                | Up-and-down jerking | Green woodpecker<br>( <i>Picus viridis</i> )                | [17]      |

|              |                          |  |      |
|--------------|--------------------------|--|------|
| Side-to-side | Head swaying             | Grey-headed woodpecker<br>( <i>Picus canus</i> )               | [17] |
|              | Head swinging            | Northern flicker<br>( <i>Colaptes auratus</i> )                | [32] |
|              | Head swinging            | White-headed woodpecker<br>( <i>Picoides albolarvatus</i> )    | [32] |
|              | Head wagging             | Williamson's sapsucker<br>( <i>Sphyrapicus thyroideus</i> )    | [32] |
|              | Head swaying             | Three-toed woodpecker<br>( <i>Picoides tridactylus</i> )       | [32] |
|              | Head turning             | Ladder-backed woodpecker<br>( <i>Picoides scalaris</i> )       | [32] |
|              | Head swinging            | Red-cockaded woodpecker<br>( <i>Picoides borealis</i> )        | [32] |
|              | Head turning             | Nuttall's woodpecker<br>( <i>Picoides nuttallii</i> )          | [32] |
|              | Head swinging            | Pileated woodpecker<br>( <i>Dryocopus pileatus</i> )           | [32] |
|              | Sway heads side to side  | Grey-headed woodpecker<br>( <i>Dendropicos spodocephalus</i> ) | [33] |
|              | Sway heads side to side  | Green woodpecker<br>( <i>Picus viridis</i> )                   | [33] |
|              | Head swaying             | Black woodpecker<br>( <i>Dryocopus martius</i> )               | [33] |
|              | Head-swaying             | Syrian woodpecker<br>( <i>Dendrocopos syriacus</i> )           | [33] |
|              | Waves bill left to right | Hairy woodpecker<br>( <i>Picoides villosus</i> )               | [94] |
|              | Waves bill left to right | Downey woodpecker<br>( <i>Picoides pubescens</i> )             | [95] |
|              | Moved from side to side  | Great spotted woodpecker<br>( <i>Dendrocopos major</i> )       | [17] |
|              | Head-swaying             | Lesser spotted woodpecker<br>( <i>Denrocopos minor</i> )       | [17] |

|           |                              |  |               |
|-----------|------------------------------|--|---------------|
| Circle    | Flying in wide circles       | Syrian woodpecker<br>( <i>Dendrocopos syriacus</i> )       | [17]          |
|           | Male circling                | Pileated woodpecker<br>( <i>Dryocopus pileatus</i> )       | [32]          |
|           | Circular flight              | Lewis' woodpecker<br>( <i>Melanerpes lewis</i> )           | [32]          |
|           | Sway head in circle          | Green woodpecker<br>( <i>Picus viridis</i> )               | [33]          |
|           | Move head in circle          | Black woodpecker<br>( <i>Dryocopus martius</i> )           | [33]          |
|           | Male circles female          | Middle spotted woodpecker<br>( <i>Dendrocopos medius</i> ) | [17]          |
|           | Circular flight path         | Lesser spotted woodpecker<br>( <i>Denrocopos minor</i> )   | [17]          |
| Spiral    | Climbing up spiraling        | Middle spotted woodpecker<br>( <i>Dendrocopos medius</i> ) | [17]          |
|           | Climb in spirals             | Grey-headed woodpecker<br>( <i>Picus canus</i> )           | [17]          |
|           | Upwards in spirals           | Green woodpecker<br>( <i>Picus viridis</i> )               | [17]          |
| Loop      | [Unavailable]                | [Unavailable]  | [Unavailable] |
| Undulate  | Undulating flight            | Yellow-bellied sapsucker<br>( <i>Sphyrapicus varius</i> )  | [32]          |
|           | Undulating flight            | Red-naped sapsucker<br>( <i>Sphyrapicus nuchalis</i> )     | [32]          |
| Figure-8  | Bill describes figure-8 path | Syrian woodpecker<br>( <i>Dendrocopos syriacus</i> )       | [17]          |
|           | Traces figure-8 course       | Green woodpecker<br>( <i>Picus viridis</i> )               | [17]          |
| U-shuttle | [Unavailable]                | [Unavailable]  | [Unavailable] |



## Appendix C

This appendix lists additional non-oscillatory motions for various birds that are best described as impulse and vibratory motions.

| Impulse and vibratory motion displays |   |           |
|---------------------------------------|---|-----------|
| Motion Description                    | Species   | Reference |
| Female dashes three to ten feet away  | Ornate tinamou<br>( <i>Nothoprocta Ornata</i> )             | [78]      |
| Erratic zig-zag flight                | Long-eared owl<br>( <i>Asio otus</i> )                      | [32]      |
| Charge, ending in a hop               | Mourning dove<br>( <i>Zenaida macroura</i> )                | [95]      |
| Short run                             | Killdeer<br>( <i>Charadrius vociferus</i> )                 | [95]      |
| Zig-zag descent                       | American woodcock<br>( <i>Scolopax minor</i> )              | [96]      |
| Erratically zigzagging                | Bald eagle<br>( <i>Haliaeetus leucocephalus</i> )           | [9]       |
| Vertical zigzags                      | Scissor-tailed flycatcher<br>( <i>Tyrannus forficatus</i> ) | [64]      |
| Frantic zig-zag courtship flight      | Cassin's kingbird<br>( <i>Tyrannus vociferans</i> )         | [32]      |
| Short leap while flapping wings       | Little bustard<br>( <i>Tetrax tetrax</i> )                  | [33]      |
| Jump up fluttering wings              | Great snipe<br>( <i>Gallinago media</i> )                   | [33]      |
| Flutter jump                          | Blue grouse<br>( <i>Dendragapus obscurus</i> )              | [32]      |
| Quivering wings                       | Northern bobwhite<br>( <i>Colinus virginianus</i> )         | [96]      |
| Flutter one wing                      | Rufous-sided towhee<br>( <i>Pipilo erythrophthalmus</i> )   | [95]      |
| Flutter wing tips                     | Chipping sparrow<br>( <i>Spizella passerina</i> )           | [95]      |
| Shaking body                          | Oldsquaw<br>( <i>Clangula hyemalis</i> )                    | [32]      |
| Head shaking                          | Gila woodpecker<br>( <i>Melanerpes uropygialis</i> )        | [32]      |
| Head shake                            | Great crested grebe<br>( <i>Podiceps cristatus</i> )        | [33]      |
| Quivering wings and tail              | Hooded crow<br>( <i>Corvus corone</i> )                     | [33]      |

## Bibliography

- [1] D. Alberts, *Talking About Mime*. Portsmouth, NH: Heinemann, 1994.
- [2] R. Arnheim, *Art and Visual Perception: A Psychology of the Creative Eye*. Berkeley: Univ. of Cal. Press, 1954.
- [3] R. Arnheim, "Visual dynamics," *American Scientist*, vol. 76, pp. 585-591, 1988.
- [4] A. O. A., Cambridge, MA.
- [5] C. Atkeson and J. Hollerbach, "Kinematic features of unrestrained vertical arm movements," *J. Neurosci.*, vol. 5, pp. 2318-2330, 1985.
- [6] D. Attenborough, *The Life of Birds*. Princeton: Princeton University Press, 1998.
- [7] M. Bastock, *Courtship: An Ethological Study*. Chicago: Aldine Publishing Company, 1967.
- [8] M. Bekoff and C. Allen, "Intentional communication and social play: how and why animals negotiate and agree to play," in *Animal Play: Evolutionary, Comparative, and Ecological Perspectives*, M. Bekoff and J. Byers, Eds. Cambridge: Cambridge Univ. Press, 1998.
- [9] J. Benyus, *Beastly Behaviors*. Reading: Addison-Wesley, 1992.
- [10] B. Bertram, *The Ostrich Communal Nesting System*. Princeton: Princeton University Press, 1992.
- [11] BioVision, Indian Wells, CA.
- [12] B. Blumberg and T. Galyean, "Multi-level direction of autonomous creatures for real-time virtual environments.," presented at SIGGRAPH, 1995.
- [13] A. Bobick, "Natural object categorization," Ph.D. Thesis, *Dept. Brain and Cognitive Science*, Massachusetts Institute of Technology, 1987.
- [14] A. Bobick, S. Intille, J. Davis, F. Baird, C. Pinhanez, L. Campbell, Y. Ivanov, A. Schutte, and A. Wilson, "The KidsRoom: a perceptually-based interactive and immersive story environment," *Presence: Teleoperators and Virtual Environments*, vol. 8, pp. 367-391, 1999.
- [15] J. Bradbury and S. Vehrencamp, *Principles of Animal Communication*. Sunderland: Sinauer Associates, 1998.
- [16] D. Brooks, E. Dunn, R. Gillmor, J. Hall-Craggs, P. Hollom, E. Nicholson, M. Ogivie, C. Roselaar, P. Sellar, K. Simmons, K. Voous, D. Wallace, and M. Wilson, *Handbook of the Birds of Europe the Middle East and North Africa*, vol. V. Oxford: Oxford University Press, 1988.

- [17] D. Brooks, E. Dunn, R. Gillmor, P. Hollom, R. Hudson, E. Nicholson, M. Ogievie, P. Olney, C. Roselaar, K. Simmons, K. Voous, D. Wallace, J. Wattel, and M. Wilson, *Handbook of the Birds of Europe the Middle East and North Africa*, vol. IV. Oxford: Oxford University Press, 1985.
- [18] R. A. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal of Robotics and Automation*, vol. RA-2, pp. 14-23, 1986.
- [19] V. Bruce, P. Green, and M. Georgeson, *Visual Perception: Physiology, Psychology, and Ecology*, Third ed. East Sussex: Psychology Press, 1996.
- [20] M. Burton, *Animal Courtship*. New York: Frederick A. Praeger, 1953.
- [21] H. Burt, *The Psychology of Birds: An Interpretation of Bird Behavior*. New York: Macmillan Company, 1967.
- [22] J. Cassell, J. Bickmore, M. Billingham, L. Campbell, K. Chang, H. Vilhjálmsón, and H. Yan, "Embodiment in conversational interfaces: Rea," presented at CHI, Pittsburgh, PA, 1999.
- [23] C. Cohen, L. Conway, and D. Koditschek, "Dynamical system representation, generation, and recognition of basic oscillatory motion gestures," presented at Int. Conf. Auto. Face and Gesture Recog., Killington, Vermont, 1996.
- [24] J. Collins and I. Stewart, "Coupled nonlinear oscillators and the symmetries of animal gaits," *J. of Nonlin. Sci.*, vol. 3, pp. 349-392, 1993.
- [25] J. Crane, "Comparative biology of Salticid spiders at Rancho Grande, Venezuela. Part IV. An analysis of display," *Zoologica*, vol. 34, pp. 159-214, 1949.
- [26] J. Crane, "Basic patterns of display in fiddler crabs (Ocypodidae, Genus uca)," *Zoologica*, vol. 42, pp. 69-82, 1957.
- [27] R. Cutler and L. Davis, "Real-time periodic motion detection, analysis, and applications," presented at IEEE Comp. Vis. Pat. Rec., Fort Collins, CO, 1999.
- [28] J. E. Cutting, D. R. Proffitt, and L. T. Kozlowski, "A biomechanical invariant for gait perception," *J. of Exp. Psych.*, vol. 4, pp. 357-372, 1978.
- [29] C. Darwin, *The Expression of the Emotions in Man and Animals*. Chicago: Univ. Chicago Press, 1872/1998.
- [30] J. Davis and A. Bobick, "The representation and recognition of human movement using temporal templates," presented at Comp. Vis. and Pat. Rec., Puerto Rico, 1997.
- [31] J. Davis and A. Bobick, "Virtual pat: a virtual personal aerobics instructor," presented at Workshop on Perceptual User Interfaces, San Francisco, CA, 1998.

- [32] P. Ehrlich, D. Dobkin, and D. Wheye, *The Birder's Handbook*. New York: Simon & Shuster Inc., 1988.
- [33] P. Ehrlich, D. Dobkin, D. Wheye, and S. Pimm, *The Birdwatcher's Handbook*. Oxford: Oxford University Press, 1994.
- [34] I. Eibl-Eibesfeldt, *Ethology: The Biology of Behavior*. New York: Holt, Rinehart and Winston, 1970.
- [35] L. Elliott and M. Read, *Common Birds and Their Songs*. Boston: Houghton Mifflin Company, 1998.
- [36] R. Estes, *The Behavior Guide to African Mammals*. Berkeley: Univ. of California Press, 1991.
- [37] R. Fagen, *Animal Play Behavior*. New York: Oxford Univ. Press, 1981.
- [38] J. Feldman, "The structure of perceptual categories," *J. of Math. Psych.*, vol. 41, pp. 145-170, 1997.
- [39] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *Journal of Neuroscience*, vol. 5, pp. 1688-1703, 1985.
- [40] L. Fleishman, "Cryptic movement in the vine snake *Oxybelis aeneus*," *Copeia*, vol. 1, pp. 242-245, 1985.
- [41] R. Fox and C. McDaniel, "The perception of biological motion by human infants," *Science*, vol. 218, pp. 486-487, 1982.
- [42] J. Freyd, "Dynamic mental representations," *Psychology Review*, vol. 94, pp. 427-438, 1987.
- [43] H. Frings and M. Frings, *Animal Communication*. Norman: Univ. of Oklahoma Press, 1977.
- [44] C. Fry and K. Fry, *Kingfishers, Bee-Eaters & Rollers*. Princeton: Princeton University Press, 1992.
- [45] J. J. Gibson, "Ecological optics," *Vision Research*, vol. 1, pp. 253-262, 1961.
- [46] J. Gould and C. Gould, *The Animal Mind*. New York: Scientific American Library, 1994.
- [47] P. Groves, "Leafy Sea Dragons," in *Scientific American*, vol. 279, 1998, pp. 85-89.
- [48] C. Harrison and A. Greensmith, *Birds of the World*. New York: Dorling Kindersley, 1993.
- [49] F. Heider and M. Simmel, "An experimental study of apparent behavior," *the American Journal of Psychology*, vol. LVII, pp. 243-259, 1944.
- [50] N. Hogan and T. Flash, "Moving gracefully: quantitative theories of motor coordination," *Trends in Neurosciences*, vol. 10, pp. 170-174, 1987.

- [51] J. Hollerbach, "An oscillation theory of handwriting," *Biological Cybernetics*, vol. 39, pp. 139-156, 1981.
- [52] M. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. IT-8, pp. 179-187, 1962.
- [53] T. Jebara and A. Pentland, "Parameterized structure from motion for 3D adaptive feedback tracking of faces," presented at IEEE Comp. Vis. Pat. Rec., San Juan, Puerto Rico, 1997.
- [54] A. Jepson and W. Richards, "A lattice framework for integrating vision modules," *IEEE Trans. Sys., Man and Cyber.*, vol. 22, pp. 1087-1096, 1992.
- [55] G. Johansson, "Visual perception of biological motion and a model for its analysis," *Perception & Psychophysics*, vol. 14, pp. 201-211, 1973.
- [56] M. Johnson, A. Wilson, B. Blumberg, C. Kline, and A. Bobick, "Sympathetic interfaces: using a plush toy to direct synthetic characters," presented at CHI, 1999.
- [57] M. Kenward, "Computing researchers take their work to the web," in *Scientific Computing World*, Oct.-Nov. 1999.
- [58] C. Kline and B. Blumberg, "The art and science of synthetic character design," presented at AISB Symposium on AI and Creativity in Entertainment and Visual Art, Edinburgh, Scotland, 1999.
- [59] J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts, "What the frog's eye tells the frog's brain," *Proceedings of the IRE*, vol. 47, pp. 1940-1951, 1959.
- [60] L. Line and F. Russell, *National Audubon Society Book of Wild Birds*. New York: Wings Books, 1997.
- [61] J. Little and J. Boyd, "Describing motion for recognition," presented at IEEE Symp. on Comp. Vis., Coral Gables, FL, 1995.
- [62] F. Liu and R. Picard, "Finding periodicity in space and time," presented at IEEE Int. Conf. Comp. Vis., Bombay, India, 1998.
- [63] D. Lowe, *Perceptual Organization and Visual Recognition*. Boston: Kluwer Academic Press, 1985.
- [64] A. E. Mace, *Ortho's Guide to the Birds Around Us*. San Francisco: Ortho Books, 1996.
- [65] Macmillan, *The Way Nature Works*. New York: Macmillan Publishing Company, 1998.
- [66] R. Mann and A. Jepson, "Non-accidental features in learning," in *Machine Learning in Computer Vision*: AAAI Press, 1994.
- [67] R. Mann, A. Jepson, and J. Siskind, "The computational perception of scene dynamics," presented at European Conference on Computer Vision, 1996.

- [68] D. Marr, "A theory of cerebellar cortex," *J. Physiol.*, vol. 292, pp. 437-470, 1969.
- [69] D. Marr, *Vision*. San Francisco: W. H. Freeman and Company, 1982.
- [70] E. P. Martins, "A comparative study of the evolution of Sceloporus push-up displays," *Am. Nat.*, vol. 142, pp. 994-1018, 1993.
- [71] G. Mather and S. West, "Recognition of animal locomotion from dynamic point-light displays," *Perception*, vol. 22, pp. 759-766, 1993.
- [72] A. Michelsen, B. Andersen, J. Storm, W. Kirchner, and M. Lindauer, "How honeybees perceive communication dance, studied by means of a mechanical model," *Behav. Ecol. Sociobiol.*, vol. 30, pp. 143-150, 1992.
- [73] P. Morasso, "Spatial control of arm movements," *Experimental Brain Research*, vol. 42, pp. 223-227, 1981.
- [74] D. Morris, "Typical intensity and its relationship to the problem of ritualization," *Behavior*, vol. II, pp. 1-12., 1957.
- [75] E. Muybridge, *Animals in Motion*. New York: Dover Publications, 1957.
- [76] National-Geographic-Video, "Heart of Africa: Forest Primeval," : Survival Anglia, 1996.
- [77] S. Niyogi and E. Adelson, "Analyzing and recognizing walking figures in XYT," presented at IEEE Comp. Vis. Pat. Rec., Seattle, WA, 1994.
- [78] A. Pearson and O. Pearson, "Natural history and breeding behavior of the Tinamou, *Nothoprocta ornata*," *Auk*, vol. 72, pp. 113-127, 1955.
- [79] R. Polana and R. Nelson, "Detection and recognition of periodic, nonrigid motion," *Int. J. Comp. Vis.*, vol. 23, pp. 261-282, 1997.
- [80] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C*, Second ed. Cambridge: Cambridge University Press, 1992.
- [81] W. Richards, *Natural Computation*, Cambridge: MIT Press, 1988.
- [82] W. Richards and A. Bobick, "Playing twenty questions with nature," in *Computational Processes in Human Vision: An Interdisciplinary Perspective*, Z. Pylyshyn, Ed. Norwood, N.J.: Ablex Publishing, 1988, pp. 3-26.
- [83] W. Richards, A. Jepson, and J. Feldman, "Priors, preferences and categorical percepts," in *Perception as Bayesian Inference*, D. Knill and W. Richards, Eds.: Cambridge University Press, 1996.
- [84] D. Rosenbaum, *Human Motor Control*. San Diego: Academic Press, Inc., 1991.
- [85] J. Rubin and W. Richards, "Boundaries of visual motion," A.I. Memo 835: Massachusetts Institute of Technology, 1985.

- [86] S. Runeson and G. Frykholm, "Visual perception of lifted weight," *J. Exp. Psych.*, vol. 7, pp. 733-740, 1981.
- [87] K. Russell and B. Blumberg, "Behavior-friendly graphics," presented at Computer Graphics International, 1999.
- [88] M. Salmon, "Waving display and sound production by Florida fiddler crabs (genus *Uca*)," *Anim. Behav.*, vol. 15, pp. 449-459, 1967.
- [89] S. Seitz and C. Dyer, "View-invariant analysis of cyclic motion," *Int. J. Comp. Vis.*, vol. 25, pp. 1-23, 1997.
- [90] K. Simmons, D. Brooks, N. Collar, E. Dunn, R. Gillmor, P. Hollom, R. Hudson, E. Nicholson, M. Ogivie, P. Olney, C. Roselaar, K. Voous, D. Wallace, J. Wattel, and M. Wilson, *Handbook of the Birds of Europe the Middle East and North Africa*, vol. III. Oxford: Oxford University Press, 1983.
- [91] K. Simmons, I. Ferguson-Lees, R. Gillmor, P. Hollom, R. Hudson, E. Nicholson, M. Ogivie, P. Olney, K. Voous, and J. Wattel, *Handbook of the Birds of Europe the Middle East and North Africa*, vol. I. Oxford: Oxford University Press, 1977.
- [92] K. Simmons, R. Gillmor, P. Hollom, R. Hudson, E. Nicholson, M. Ogivie, P. Olney, C. Roselaar, K. Voous, D. Wallace, and J. Wattel, *Handbook of the Birds of Europe the Middle East and North Africa*, vol. II. Oxford: Oxford University Press, 1980.
- [93] R. Smith, "Paternity assurance and altered roles in the mating behaviour of a giant water bug, *Abedus herberti* (Heteroptera: Belostomatidae)," *Animal Behavior*, vol. 27, pp. 716-725, 1979.
- [94] D. Stokes, *A Guide to Bird Behavior*, vol. 1. Boston: Little, Brown and Company, 1979.
- [95] D. Stokes and L. Stokes, *A Guide to Bird Behavior*, vol. 2. Boston: Little, Brown and Company, 1983.
- [96] D. Stokes and L. Stokes, *A Guide to Bird Behavior*, vol. 3. Boston: Little, Brown and Company, 1989.
- [97] D. Stokes and L. Stokes, *The Hummingbird Book*. Boston: Little, Brown and Company, 1989.
- [98] D. Terzopoulos, X. Tu, and R. Grzeszczuk, "Artificial fishes: autonomous locomotion, perception, behavior, and learning in a simulated physical world," *Artificial Life*, vol. 1, pp. 327-351, 1994.
- [99] N. Tinbergen, *The Study of Instinct*. London: Oxford Univ. Press, 1951.
- [100] N. Tinbergen, "Derived activities: their causation, biological significance, origin, and emancipation during evolution," *Quart. Rev. Biol.*, vol. 27, pp. 1-32, 1952.

- [101] N. Tinbergen, *The Animal in Its World*, vol. 1. Cambridge: Harvard University Press, 1972.
- [102] E. Todorov, "Studies of goal directed movements," Ph.D. Thesis, *Dept. Brain and Cognitive Sciences*, Massachusetts Institute of Technology, 1998.
- [103] P.-S. Tsai, K. Keiter, T. Kasparis, and M. Shah, "Cyclic motion detection," *Pat. Rec.*, vol. 27, pp. 1591-1603, 1994.
- [104] E. Turner, "Social feeding in birds," *Behaviour*, vol. 24, pp. 1-46, 1965.
- [105] T. Tynning, *A Guide to Amphibians and Reptiles*. Boston: Little, Brown and Company, 1990.
- [106] M. Vernon, *The Psychology of Perception*. Baltimore: Penguin Books, 1962.
- [107] P. Viviani and T. Flash, "Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 21, pp. 32-53, 1995.
- [108] K. von Frisch, *The Dance Language and Orientation of Bees*. Cambridge: Harvard Univ. Press, 1967.
- [109] A. P. Witkin and J. M. Tenenbaum, "On the role of structure in vision," in *Human and Machine Vision*, J. Beck, B. Hope, and A. Rosenfeld, Eds. New York: Academic Press, 1983, pp. 481-543.
- [110] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real-Time Tracking of the Human Body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 780-785, 1997.