

DA  
2074 (HG)  
1997

**Gesture Recognition by Motion Extraction and  
Multivariate Analysis**

モーション特徴抽出と多変量解析に基づく  
ジェスチャ認識

**Doctoral Program in Engineering  
University of Tsukuba**

**March 2000**

**Bisser Roumenov Raytchev**

寄贈  
ライチエフ  
ビセル氏

00301862

# **Gesture Recognition by Motion Extraction and Multivariate Analysis**

## **Abstract**

Bisser Roumenov Raytchev  
Tsukuba University  
March 2000

This thesis proposes a user-independent gesture recognition system, which operates in real-time under real-world conditions, and as such could be useful as a building block in a more sophisticated human-computer interface. In the proposed system, efficient and robust extraction/representation of motion patterns is achieved by utilizing *relative-motion dependent primitive features*. This bottom-up approach for primitive feature extraction in the early processing stage of the system is combined with *multivariate analysis based top-down learning* and *Dynamic Buffer Structures* for online gesture segmentation. The proposed system exhibits good generalization abilities even if trained with only a few samples. In test experiments conducted on several different data sets it was observed that system's performance is robust to changes in the background and in the illumination conditions, to subjects' body size and external appearance, able to cope with the non-uniformity in the performance speed of the gestures, etc., all of which are important conditions for a successful application in human-computer interfaces. Gesture recognition in the present system is not restricted to hand gestures only – larger-scale whole body gestures can be processed simultaneously with relatively smaller-scale hand or head gestures. No manual segmentation of any kind or use of markers, sensors, etc. is necessary. Also, since no domain knowledge is utilized, the system proposed here can be easily adapted and applied to other problems involving motion recognition.

## Acknowledgments

First, I would like to thank my advisor, Dr. Nobuyuki Otsu, Director of Machine Understanding Division, ETL, and concurrently Professor of the post graduate school of Tsukuba University, for his consistent support, guidance and encouragement throughout the last three years during which this research was conducted. I would like also to thank Dr. K. Sakaue for accepting me in the Adaptive Vision Laboratory at ETL, and for providing a cooperative and productive work environment. Thanks are also due to all members of the Adaptive Vision Lab and Infodynamics Lab at ETL for many helpful discussions, technical assistance and creative cooperation. Especially I would like to thank Dr. O. Hasegawa and Dr. T. Kurita for their help respectively at the early and later stages of this research.

Also I would like to thank to the thesis committee members from Tsukuba University, Prof. S. Itabashi, Prof. K. Toraichi, Prof. Y. Ohta and Prof. K. Sakaue for their advice and encouragement.

Last but not least many thanks to my wife, my parents, my teachers, my sponsors and my friends for their constant support and encouragement.

# CONTENTS

<b>List of Figures</b>	v
<b>1. Introduction</b>	<b>1</b>
1.1. The subject of this thesis .....	1
1.2. Motion recognition applications .....	2
1.3. Outline of the approach .....	4
<b>2. Related Work</b>	<b>7</b>
1.1. Methods for motion extraction .....	7
1.2. The study of gestures .....	9
1.3. Motion-based gesture recognition .....	11
<b>3. Primitive features for relative motion dependent feature extraction</b>	<b>14</b>
3.1. Psychophysical evidence in support of the current approach .....	16
3.2. Primitive features for relative motion dependent feature extraction .....	24
3.3. Discussion and further extensions of the approach .....	37
<b>4. Multivariate analysis based learning</b>	<b>44</b>
4.1. Motivation for the approach .....	44
4.2. Projection of primitive feature space to discriminant feature space .....	46
4.3. Output from the classifier .....	49
4.4. Combining ensembles of classifiers .....	52

<b>5. Online gesture segmentation by Dynamic Buffer Structures</b>	<b>56</b>
5.1. Gesture segmentation problems .....	56
5.2. Dynamic Buffer Structures for gesture segmentation .....	59
5.3. Parameter determination and discussion .....	63
<b>6. Experimental results</b>	<b>66</b>
6.1. Multimodal database of gestures with speech (MMDB) .....	66
6.2. "Real-world" data .....	68
6.2.1. Single user .....	69
6.2.2. Different subjects .....	70
6.2.3. Different backgrounds and illumination conditions .....	73
6.2.4. Tests for size invariance .....	74
6.2.5. Tests under some more extreme conditions .....	76
<b>7. Conclusion and further work</b>	<b>79</b>
<b>BIBLIOGRAPHY</b>	<b>82</b>
<b>LIST OF PUBLICATIONS</b>	<b>88</b>