LNAI 2915

Antonio Camurri
Gualtiero Volpe (Eds.)

Gesture-Based Communication in Human-Computer Interaction

5th International Gesture Workshop, GW 2003
Genova, Italy, April 2003
Selected Revised Papers

Springer

Antonio Camurri Gualtiero Volpe (Eds.)

Gesture-Based Communication in Human-Computer Interaction

5th International Gesture Workshop, GW 2003
Genova, Italy, April 15-17, 2003
Selected Revised Papers

Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Antonio Camurri
Gualtiero Volpe
University of Genova
DIST, InfoMus Lab
Viale Causa 13, 16145 Genova, Italy
E-mail: music@dist.unige.it, volpe@infomus.dist.unige.it

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

CR Subject Classification (1998): I.2, I.3.7, I.5, I.4, H.5.2

ISSN 0302-9743
ISBN 3-540-21072-5 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Olgun Computergrafik
Printed on acid-free paper SPIN: 10988329 06/3142 5 4 3 2 1 0

Preface

Research on the multifaceted aspects of modeling, analysis, and synthesis of human gesture is receiving growing interest from both the academic and industrial communities. On one hand, recent scientific developments on cognition, on affect/emotion, on multimodal interfaces, and on multimedia have opened new perspectives on the integration of more sophisticated models of gesture in computer systems. On the other hand, the consolidation of new technologies enabling "disappearing" computers and (multimodal) interfaces to be integrated into the natural environments of users are making it realistic to consider tackling the complex meaning and subtleties of human gesture in multimedia systems, enabling a deeper, user-centered, enhanced physical participation and experience in the human-machine interaction process.

The research programs supported by the European Commission and several national institutions and governments individuated in recent years strategic fields strictly concerned with gesture research. For example, the DG Information Society of the European Commission (www.cordis.lu/ist) supports several initiatives, such as the "Disappearing Computer" and "Presence" EU-IST FET (Future and Emerging Technologies), the IST program "Interfaces & Enhanced Audio-Visual Services" (see for example the project MEGA, Multisensory Expressive Gesture Applications, www.megaproject.org), and the IST strategic objective "Multimodal Interfaces." Several EC projects and other funded research are represented in the chapters of this book.

A wide range of applications can benefit from advances in research on gesture, from consolidated areas such as surveillance to new or emerging fields such as therapy and rehabilitation, home consumer goods, entertainment, and audiovisual, cultural and artistic applications, just to mention only a few of them.

This book is a selection of revised papers presented at the Gesture Workshop 2003, the 5th International Workshop on Gesture and Sign Language-Based Human-Computer Interaction, held in Genoa, Italy, during April 15–17, 2003.

The International Gesture Workshop is a forum where researchers working on gesture-based interfaces and gestural interaction present and exchange ideas and research currently in progress, with a crossdisciplinary focus. GW2003 was the fifth workshop after the 1996 Gesture Workshop in York (UK), considered as the starting event. Thenceforth, International Gesture Workshops have been held roughly every second year, with fully reviewed postproceedings typically published by Springer-Verlag.

As an indicator of the continuously growing interest of the scientific community in gesture-mediated human-computer interaction and human-language technology, a large number of high-quality submissions was received. The program included invited talks, oral presentations of long and short papers, presentations of posters, and demonstrations: around 90 contributors from 20 different countries offered a broad overview of the state of the art in many research fields related to gesture-based communication. Over 170 delegates attended the workshop.

This workshop was organized by the InfoMus Lab at the DIST, University of Genoa, and was supported by the aforementioned EC IST MEGA project and by the Opera House Teatro Carlo Felice of Genova. We wish to thank Gennaro Di Benedetto, Sovrintendente of the Teatro dell'Opera Carlo Felice and his staff (with particular thanks to Rita Castello, Graziella Rapallo and Giampaolo Sperini), APT Genova and Agenzia Regionale per la Promozione Turistica della Liguria, the invited speakers Frank Pollick (Department of Psychology, University of Glasgow, UK) and Shuji Hashimoto (Department of Applied Physics, Waseda University, Tokyo, Japan), the Scientific Committee, the session chairs, Barbara Mazzarino and the other members of the Local Organizing Committee (Roberto Chiarvetto, Roberto Dillon and Paolo Coletta), the staff and the students of the DIST InfoMus Lab who helped in the organization, and all the presenters and attendees. We thank also Martino Musso (Lever) for the support to the organization of the event, Eidomedia and NumenSoft.

November 2003

Antonio Camurri
Gualtiero Volpe

International Programme Committee and Book Reviewers

Daniel Arfib	CNRS-LMA, Marseille, France
Nicola Bernardini	Media Innovation Unit, Firenze Tecnologia, Italy
Annelies Braffort	LIMSI-CNRS, Université de Paris Sud, France
Roberto Bresin	KTH, Stockholm, Sweden
Antonio Camurri	InfoMus Lab, DIST, University of Genoa, Italy
Giovanni De Poli	CSC, DEI, University of Padua, Italy
Anders Friberg	KTH, Stockholm, Sweden
Sylvie Gibet	Valoria, Université de Bretagne Sud, France
Rolf Inge Godøy	Section for Musicology, University of Oslo, Norway
Marianne Gullberg	Max-Planck Institute of Psycholinguistics, Nijmegen, The Netherlands
Shuji Hashimoto	Waseda University, Tokyo, Japan
Richard Kennaway	University of East Anglia, Norwich, UK
Marc Leman	IPEM, Ghent University, Belgium
Pierre-François Marteau	Valoria, Université de Bretagne Sud, France
Gildas Menier	Valoria, Université de Bretagne Sud, France
Catherine Pelachaud	LINC, Paragraphe, IUT of Montreuil, University of Paris 8, France
Isabella Poggi	Dipartimento di Scienze dell'Educazione, University Roma Tre, Italy
Frank E. Pollick	Department of Psychology, Glasgow University, UK
Timo Sowa	Artificial Intelligence Group, Faculty of Technology, University of Bielefeld, Germany
Johan Sundberg	KTH, Stockholm, Sweden
Daniel Thalman	EPFL, Lausanne, Switzerland
Leon van Noorden	European Commission, DG Information Society
Gualtiero Volpe	InfoMus Lab, DIST, University of Genoa, Italy
Ipke Wachsmuth	Artificial Intelligence Group, Faculty of Technology, University of Bielefeld, Germany
Marcelo Wanderley	Faculty of Music, McGill University, Canada

Table of Contents

Foundational Issues

Gesture Tracking

Gesture Recognition

Gesture Notation and Synthesis

Multimodal Gestural Interfaces

Gesture Analysis: Invariant Laws in Movement

Sylvie Gibet, Jean-François Kamp, and Franck Poirier

Valoria, Université de Bretagne Sud, Campus de Tohannic, rue Yves Mainguy
F-56000 Vannes, France
{Sylvie.Gibet,Jean-Francois.Kamp,Franck.Poirier}@univ-ubs.fr

Abstract. This paper presents gesture analysis under the scope of motor control theory. Following the motor program view, some studies have revealed a number of invariant features that characterize movement trajectories in human hand-arm gestures. These features express general spatio-temporal laws underlying coordination and motor control processes. Some typical invariants are described and illustrated for planar pointing and tracing gestures. We finally discuss how these invariant laws can be used for motion edition and generation.

1 Introduction

With the massive development of Human-Computer Interaction (HCI), new systems try to take advantage of the expressive power of gestures. At first, gesture interaction has been reduced to simple command interfaces. More recently, techniques such as capturing body movements, recognizing and interpreting human actions, and animating virtual humans, have given rise to a number of virtual reality applications with more natural interfaces. In such applications, the produced gestures should be perceived as natural and should follow biomechanical or psychomotor laws characterizing human movement.

This paper presents gesture analysis results under the scope of motor control theory. The analysis leads to the identification of significant parameters that characterize some classes of gesture (simple pointing, point-to-point and rhythmic gestures), extracted from real data in the Cartesian space. The main issues are regarded as a way to understand the mechanisms underlying the control and coordination of movement, both in terms of general principles involving central processes of task planning and in terms of peripheral processes emerging from the bio-mechanical system. We examine the principles that can be expressed as invariant features in movement trajectories, the most well-known ones being the Fitt's law and the Two-Third Power law. The overview is restricted to planar pointing and tracing movements, with patterns which are not previously learnt. We don't intend to demonstrate that these laws prefigure organizational principles explicitly used by the Central Nervous System (CNS) to produce movement, but we give some elements that can be used for motion edition or generation in Human Machine interaction or computer animation.

A. Camurri and G. Volpe (Eds.): GW 2003, LNAI 2915, pp. 1–9, 2004.

2 Motor Control Theories

During the last decades, two main classes of control theories have been developed. The first one, related to the motor program concept, assumes that there exists a central instance of control which exploits an internal representation of motion to determine the appropriate motor control. The second class postulates that the processes responsible for motor control can be found in the organization of the biomechanical system itself rather than in imposed laws.

2.1 Motor Program

The motor program concept has been introduced by the cognitive science community [1]. It can be defined as the ability to build a set of organized commands in space and time before performing gesture. This notion support the hypothesis that there is no sensorial feedback during the execution of motion. For instance, goal-based gestures such as pointing gestures highlight some kind of pre-programmed activity, the corresponding motion being executed without visual or proprioceptive feedback. To take into account longer motion and also parametrical features, this notion of motor program has been replaced by Generalized Motor Programs (GMP) [2], [3], referring to a set of generic rules associated to specific motion classes. These rules could be issued from a learning process, mapping the efferent command, the sensory feedback, the environmental context and the result of the action. GMP partially explain invariant properties characterizing motion which could be coded in general schemes, the adaptation to the performing context being realized by the instantiation of specific parameters. The search for the invariant properties has given rise to the setting of motion laws which are presented in section 3.

2.2 Biomechanical Properties of the Effectors

Other approaches suggest that limb dynamics and biomechanical properties can contribute significantly to motion control. Within this point of view, the dynamical system paradigm considers that the control processes responsible for the formation of trajectories are not explicitly programmed, but are emerging properties of the dynamics of the system itself.

According to the equilibrium point hypothesis, movements arise from shifts in the equilibrium position of the muscles. The motor program specifies a succession of discrete equilibrium points [4]. Between two equilibrium points (targets), motion is generated according to the dynamics of the mechanical system. Several equilibrium point models have been proposed. The most weel-known is the Bizzi model [5], inspired from the Feldman model. These models are essentially exploited for discrete multi-point tasks, or for elliptical trajectories generated from the specification of via-points under isometric conditions. After observing that discontinuous displacement of equilibrium points did not lead to the production of realistic motion, Hogan suggested that these equilibrium points follow a virtual trajectory, thus forcing the muscular forces to satisfy this constraint [6].

Numerous research studies have been carried out for cyclic or rhythmic movements. The concept of Motor Pattern Generation MPG has been developed for respiratory movements or locomotion [7-8]. These MPG explain the activity of neuronal circuits easily identifiable.

Other theories lying on nonlinear dynamical systems try to show the importance of accounting for physical properties in generating motion. They have given rise to models of coordination structures including a set of nonlinear oscillating systems mutually coupled. But as opposed to MPG models, the oscillators are not directly associated to specific neural structures [9-10].

3 Invariant Laws

Despite the great variability of gestures, some invariant features of the motor performance have been highlighted in the past years. The enhanced hypothesis is that these spatio-temporal invariants express general laws underlying the organization of motricity. Several kinematic laws characterizing human volunteer movements have been proposed by cognitive scientists or neurophysiologists. These laws are satisfied for a large variety of gestures, including goal-directed gestures such as simple or multiple pointing gestures, but also repetitive cyclic gestures in two or three dimensions. Without trying to give an exhaustive view of these laws, we present in this paper the most typical invariants of movement trajectories. They include invariance of velocity profile, Fitts' law showing a relationship between kinematics and geometrical properties of the trajectory for pointing movements, two-third power law for rhythmic tasks and minimum jerk expressing smoothness of trajectories.

The presented laws have been highlighted from real signals captured with a pen on a graphical WACOM A3 tablet.

3.1 Invariance of the Velocity Profile

Multi-point movements produce velocity profiles whose global shape is approximately bell-shaped. Moreover, this shape presents an asymmetry which depends on the speed of the movement. As the speed increases the curve becomes more symmetrical until the direction of the asymmetry is reversed [11-13]. This law is shown in the figure 1 for pointing gestures with different speeds and distances to the target.

3.2 Isochrony Principle and Fitts' Law

The principle, originally established by Freeman [14], and known as the Isochrony principle, expresses the invariance of the execution duration of a movement in relation to its amplitude. It has been observed that there exists a spontaneous tendency to increase the velocity of the motion according to the distance to run, when no constraint on the mean velocity is imposed.

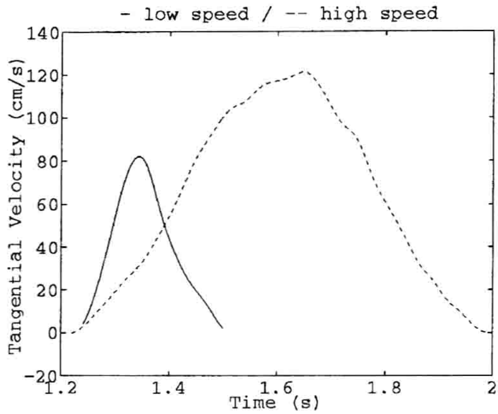

Fig. 1. Two velocity profiles: low speed (max ≈ 80 cm/s ; distance to target ≈ 10.3 cm) and high speed (max ≈ 120 cm/s ; distance to target ≈ 46.6 cm) for pointing gestures.

In this line of research, the Fitt's law has quantified this constancy in time duration for rapid movements between targets in a plane. First experiments, realized by Woodworth [15] and then by Fitts [16] consisted in executing back-and-forth movements with a pen, the distance between the targets and the width of the targets being varied over experiences. The Fitts' law (1) that can be referred to the task's index of difficulty I_d, expresses the relation between the duration of motion T, the amplitude required A and the target width W.

$$T = a + b.I_d = a + b.\log_2\left(\frac{2.A}{W}\right) \quad (1)$$

$$T = a + b.\log_2\left(\frac{A}{W} + c\right) \qquad c = ½ \text{ [17] or } c = 1 \text{ [18]} \quad (2)$$

where a and b are constants determined in an empirical way.

This law is illustrated in the figure 2, for simple pointing gestures. In these experiments, the distance to the target varies from 10.3 cm to 38.4 cm and the target width W is constant.

The Fitt's law can be verified in diversified contexts for pointing gestures requiring a good accuracy of the hand endpoint. It is extensively used in its original form (1) or in variant forms (2), in particular to investigate the relative performance of devices used in Human-Machine interfaces. It can be extended to other classes of motion, as for example multi-scaled pointing gestures [19], [20].

3.3 Two-Third Power Law

For handwriting and drawing movements performed in a plane, studies by Viviani and Terzuolo shown that there is a relationship between the kinematics of elliptical

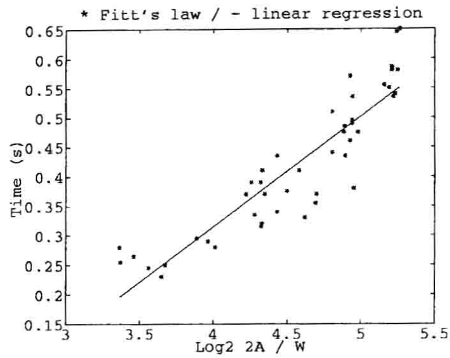

Fig. 2. Fitts' law: 45 pointing measurements (W = 2 cm and A = 10.3 cm … 38.4 cm). Linear regression of Time against $\log_2 2A/W$ gives a slope of 0.19. Correlation coefficient between measurements and linear regression = 0.9.

motion and the geometrical properties of the trajectory. This has given rise to the so-called Two-Third Power law [21-22] that establishes a relation between the angular velocity ω and the curvature C of the end-effector trajectory (3).

$$\omega(t) = kC(t)^{\frac{2}{3}} \tag{3}$$

Or equivalently :

$$v(t) = k\, R(t)^{\beta} \quad where\ \beta = \frac{1}{3} \tag{4}$$

$$v(t) = \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} \tag{5}$$

$$R(t) = \frac{v(t)^3}{\left|\dot{x}(t).\ddot{y}(t) - \ddot{x}(t).\dot{y}(t)\right|} \tag{6}$$

where $v(t)$ is the tangential velocity and $R(t)$ the radius of curvature.

A more recent formulation of this law extends the validity of the original law, for a wider class of movements [23]:

$$v(t) = K\left(\frac{R(t)}{1+\alpha R(t)}\right)^{\beta} \quad \text{with } 0<\alpha<1 \text{ and } K>0 \tag{7}$$

The exponent β takes values close to 1/3 (for adults). Figure 3 illustrates the power law for human elliptical movements.

The 2/3 power law has been suggested as a fundamental principle that the CNS may employ to constrain trajectory formation. Viviani supposes that complex movement can be decomposed into units of motor action. He suggests that the portions of movement over which the factor K is approximately constant correspond to chunks of motor planning. Thus, the velocity gain factor K can be used as a parameter to segment complex movements. In the case of tracing movements of ellipses with a given size and speed, K is constant during all the pattern ; its value depends among other factors on the frequency of the movement and on the eccentricity of the ellipse. For more complex graphical patterns with several loops, it is possible to identify several units corresponding to these loops.

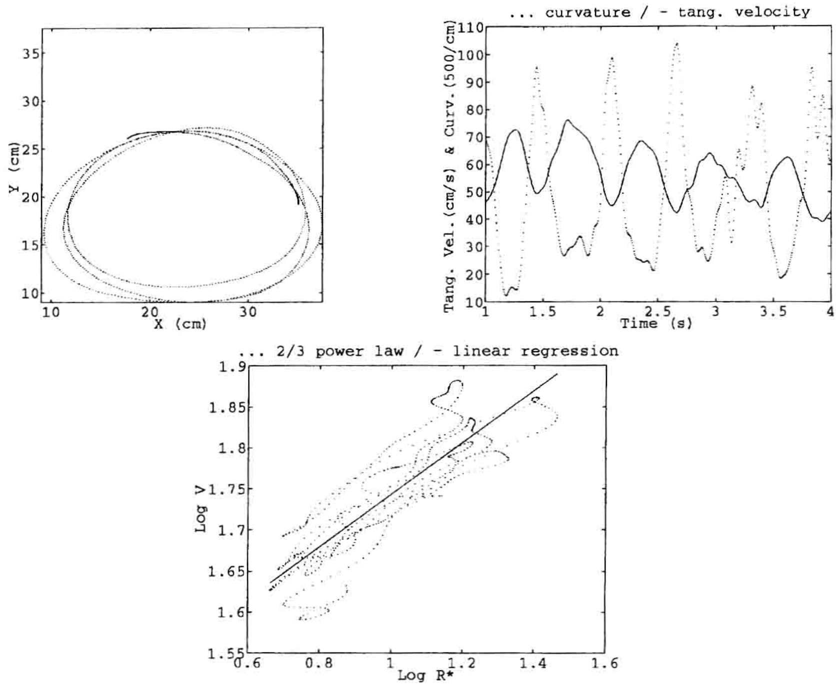

Fig. 3. Top left: human elliptical movement in the plane. Top right: tangential velocity and curvature displayed on the same scale. Bottom: log of tangential velocity versus log of R* where R* = R / (1+ α R), α = 0.01. Linear regression of log V against log R* gives a slope of 0.32 ≈ 1/3.

Supporting evidence for the power law has been provided in a large variety of studies. However, the law is most of the time obeyed for planar drawing patterns of relatively small size.

3.4 Smoothness Considerations

The Minimum Jerk model proposes another interpretation to motor production, providing a solution to the problem of trajectory formation. For point-to-point movements, it ensures that among all possible solutions for generating a trajectory, the motor control system chooses the one that maximizes the smoothness of the movement. This solution complies with the minimization of a global cost function C, ex-