Sylvie Gibet
Nicolas Courty
Jean-François Kamp (Eds.)

LNAI 3881

Gesture in Human-Computer Interaction and Simulation

6th International Gesture Workshop, GW 2005
Berder Island, France, May 2005
Revised Selected Papers

Springer

Sylvie Gibet Nicolas Courty
Jean-François Kamp (Eds.)

Gesture in Human-Computer Interaction and Simulation

6th International Gesture Workshop, GW 2005
Berder Island, France, May 18-20, 2005
Revised Selected Papers

Springer

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

Volume Editors
Sylvie Gibet
Nicolas Courty
Jean-François Kamp
Université de Bretagne Sud
Laboratoire VALORIA, Centre de Recherche Yves Coppens
Campus de Tohannic, rue Yves Mainguy, 56000 Vannes, France,
E-mail: {Sylvie.Gibet,Nicolas.Courty,Jean-Francois.Kamp}@univ-ubs.fr

Library of Congress Control Number: 2006920788

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

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743
ISBN-10 3-540-32624-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-32624-3 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11678816 06/3142 5 4 3 2 1 0

Preface

The international Gesture Workshops have become the leading interdisciplinary events for dissemination of the latest results on gesture-based communication. The goal of these workshops is to bring together researchers who want to meet and share ideas on advanced research on gesture related to multidisciplinary scientific fields. Depending on the fields, the objectives can be very different. While physiology and biomechanics aim to extract fundamental knowledge of physical gesture, computer sciences try to capture different aspects of gesture and extract features that help to identify, interpret or rebuild the underlying mechanisms of communication gestures. Other approaches and methodologies are followed by cognitive sciences and linguistics, bringing a complementary understanding of motor control and gesture meaning. The results can be enhanced by technological applications or demonstrations. For example, gestural interaction in an augmented or virtual reality context leads to active application areas. Since 1996 gesture workshops have been held approximately every two years, with full post-proceedings usually published by Springer.

Gesture Workshop 2005 (GW 2005) was organized by VALORIA, at the University of Bretagne Sud (Vannes, France), and was held on Berder Island, Morbihan (France) during May 18-20, 2005. This event, the sixth in a highly successful workshop series, was attended by more than 70 participants from all over the world (13 countries). Like the previous events, GW 2005 aimed to encourage multidisciplinary exchanges by providing an opportunity for participants to share new results, show live demonstrations of their work, and discuss emerging directions on topics broadly covering the different aspects of gesture. The very special area where the workshop took place (a small island in the Gulf of Morbihan) provided an occasion for lively discussions and establishment of future collaboration on research centered on gesture as a means of communication. A large number of high-quality submissions was received, which made GW 2005 a great event for both industrial and research communities interested in gesture-based models relevant to human–computer interaction and simulation.

This book is a selection of revised papers presented at Gesture Workshop 2005. Containing 24 long papers and 14 short papers, it offers a wide overview of the most recent results and work in progress related to gesture-based communication. Two contributions on major topics of interest are included from two invited speakers. The contribution from Jean-Louis Vercher (Movement and Perception Lab., Marseille, France) is concerned with fundamental issues of biological motion, and their link with the perception and the synthesis of realistic motion. The contribution from Ronan Boulic et al. (EPFL, Switzerland) highlights the potential of some well-known computer animation methods for motion synthesis. The book covers eight sections of reviewed papers relative to the following themes:

- Human perception and production of gesture
- Sign language representation
- Sign language recognition
- Vision-based gesture recognition
- Gesture analysis
- Gesture synthesis
- Gesture and music
- Gesture interaction in multimodal systems

Under the focus of gesture in Human-Computer Interaction and Simulation, the book encompasses all aspects of gesture studies in emerging research fields. Two sections are devoted to sign language representation and recognition. Pertinent features extracted from captured gestures (signal, image) are used for processing, segmentation, recognition or synthesis of gestures. These topics concern at least three sections of the book. Different kinds of applications are considered, including for example expressive conversational agents, gesture interaction in multimodal systems, and gesture for music and performing arts.

The workshop was supported by the University of Bretagne Sud (France), the French Ministry of Research, the *Conseil Régional de Bretagne* and the *Conseil Général du Morbihan*: we are very grateful for their generous financial support. GW 2005 also received some financial support from COST-European Science Foundation. In particular, the Cost287-ConGAS action, mainly concerned with Gesture Controlled Audio Systems, was strongly represented within the workshop, and we are grateful to the delegates for their contribution to the event and the book. Thanks also to France Telecom R&D (a French telecommunication society) which generously contributed to the sponsoring of GW 2005, and participated in the forum by presenting very relevant demonstrations.

We would also like to express our thanks to the local Organizing Committee (Sylviane Boisadan, Alexis Héloir, Gildas Ménier, Elisabeth Le Saux, Joël Révault, Pierre-François Marteau) as well as Gersan Moguérou for webmastering the GW2005 Internet site. We are also grateful to the university staff and the PhD students from VALORIA who helped in the organization of the workshop.

Finally, the editors are thankful to the authors of the papers, as well as the international reviewers. As a result of their work, this volume will serve as an up-to-date reference for researchers in all the related disciplines.

December 2005

Sylvie Gibet
Nicolas Courty
Jean-François Kamp

Reviewers

Bruno Arnaldi	IRISA, France
Ronan Boulic	Virtual Reality Lab., EPFL, Switzerland
Annelies Braffort	LIMSI-CNRS, France
Antonio Camurri	InfoMus Lab., DIST, University of Genova, Italy
Nicolas Courty	Valoria, University of Bretagne Sud, France
Winand Dittrich	University of Hertfordshire, UK
Sylvie Gibet	Valoria, University of Bretagne Sud, France
Philippe Gorce	LESP, University of Toulon et du Var, France
Marianne Gullberg	Max Planck Institute, The Netherlands
Alexis Héloir	Valoria, University of Bretagne Sud, France
Jean-François Kamp	Valoria, University of Bretagne Sud, France
Richard Kennaway	Norwich, UK
Stefan Kopp	Bielefeld University, Germany
Seong-Whan Lee	Korea University, Korea
Pierre-François Marteau	Valoria, University of Bretagne Sud, France
Gildas Ménier	Valoria, University of Bretagne Sud, France
Franck Multon	LPBEM, University of Rennes 2, France
Hermann Ney	Aachen University, Germany
Catherine Pelachaud	University of Paris 8, France
Danielle Pelé	France Telecom R&D, France
Timo Sowa	Bielefeld University, Germany
Jean-Louis Vercher	LMP, University of the Mediterranean, Marseille, France
Christian Vogler	Washington University, USA
Gualterio Volpe	InfoMus Lab., DIST, University of Genova, Italy
Ipke Wachsmuth	Bielefeld University, Germany
Marcelo Wanderley	McGill University, Canada

Table of Contents

Sign Language Recognition

Long Papers

Short Papers

Vision-Based Gesture Recognition

Long Papers

Short Papers

Gesture Analysis

Long Papers

Short Papers

Gesture Synthesis

Invited Paper

Long Papers

Short Papers

Gesture and Music

Long Papers

Perception and Synthesis of Biologically Plausible Motion: From Human Physiology to Virtual Reality

Jean-Louis Vercher

UMR CNRS 6152 Mouvement et Perception, Institut Fédératif de Recherche
Etienne-Jules Marey, Université de la Méditerranée Marseille, France
vercher@laps.univ-mrs.fr

Abstract. To model and simulate human gesture is a challenge which takes benefit from a close collaboration between scientists from several fields: psychology, physiology, biomechanics, cognitive and computer sciences, etc. As an a priori requirement, we need to better understand the so-called laws of biological motions, established all along the 20^{th} century. When modelled and used to animate artificial creature, these laws makes these creatures (either virtual or robotic) move in a much more realistic, life-like, fashion.

1 Introduction

A virtual reality (VR) system is expected to provide realistic representations of objects. The realistic character applies at least as much to the behaviour of the objects as to their aspect [1]. A moving object must particularly comply with certain rules: at first, the laws of physics of course, and in particular those of the Galilean kinematics and Newtonian dynamics. The compliance with these rules confers realistic properties to the environment (i.e. gravity) as to the objects (inertia, surface properties, constitution, etc). In the particular case where the animated object corresponds to a living being (animal, human) or supposed such, additional rules are essential, to obtain that the simulated item is perceived as being alive. Indeed, many studies in the field of psychology of perception revealed the existence of biological "signatures" in the movement of the living beings. The existence of these signatures in representations of moving objects, not only are enough "to animate" (within the meaning of "giving life to") the objects, but are essential to allow their recognition as products of a biological activity. All along this paper, we will review the principal studies related to the perception of biological motion, we will see how the designers of virtual reality applications, as those of multi-media and cinema industries take advantage from this knowledge, and we will see finally how virtual reality can help perception psychology to better understand the phenomena which gives the character ALIVE to animated symbolic systems.

2 The Perception of Biological Motion

Studies in psychology and neurophysiology obviously show that the movements of a human body can be easily recognized and identified in their biological, living nature.

S. Gibet, N. Courty, and J.-F. Kamp (Eds.): GW 2005, LNAI 3881, pp. 1 – 12, 2006.

The literature on the perception of the so-called biological motion is abundant and varied. This literature really finds its source in the seminal work of Johansson [2], who characterized biological motion as referring to the ambulatory patterns of bipeds and terrestrial quadrupeds.

2.1 Characteristic Points Are Worth a Complex Picture: Johansson

Johansson [2] filmed in total darkness walkers with, as only visible elements, some lights attached to the main joints of the body (Fig. 1). He showed that such points, moving on a uniform background, were perceived by observers as indicating human movements, in the absence of any other visual index. At the end of the 19th century, Muybridge and Marey had already, in an implicit way, used this faculty of our brain to reconstitute the complex gesture from a finished number of points (connected or not between them by straight lines).

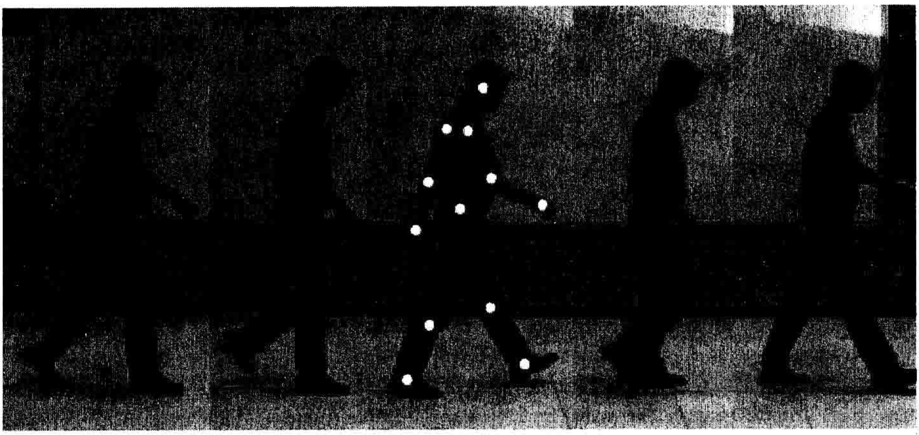

Fig. 1. Example of a walker on which markers are placed on specific joint

Johansson's experiment has been frequently reproduced and confirmed, and has become a traditional chapter of perception psychology. When the filmed people dance, the observer can eventually identify the type of dance [3], and to distinguish the men from the women [4]. This capability to perceive the human or biological character from sparse indices is innate or almost: it has been observed in children as early as 4 to 6 months after birth [5]. The biological nature of movement may be identified even when the group of dots corresponds to the movement of an animal [6], it thus does not act of a specific effect. It even seems that we have the capability to infer the shape and the physical properties of the body from a simple group of moving dots moving. Many visual demonstrations of this effect can be observed online [7, 8].

Most of the studies related to the perception of biological motion do not carry in them any attempt at explanation: they do nothing but show-up the existence of the phenomenon, and in particular they highlight the need for maintaining the characteristics kinematics of the points, whatever they are. Some attempts at modelling of this perceptive capacity were however proposed, in order to explain how the three-dimensional structure of the movement of the limbs of an animal could be calculated starting from the two-dimensional movements of some markers projected in the plan of the image [9]. At least two of these attempts deserve to be known.

2.2 Paolo Viviani and the Motor Theory of Perception

The determination of invariant characteristics of movement constitutes a millstone of crucial importance to the understanding of the fundamental principles of organization of biological motor control, concerning in particular the role of the central nervous system (CNS). Although his work was not directly related to the problems of the perception of biological motion (but rather to automatic signature recognition), Viviani contributed to the comprehension of the phenomenon. Seeking to identify an invariant in the morphogenesis of writing, he demonstrated the existence of a non-linear relation linking the angular velocity of the hand to the trajectory curvature [10, 11]:

$$a(t)=kc(t)^{2/3} \quad (1)$$

This relation, extremely robust (the human gesture cannot violate it) strongly conditions our perception. Its non-observance leads an observer to confuse the shape of the trajectory: a circle becomes an ellipse and vice and versa. When for example one observes a luminous point moving on an elliptical path according to kinematics corresponding to a circle (constant angular speed), the observer perceives the point as moving along a circle. The vision is not the only sensorial modality concerned: passive movements of the hand induced by a computer via a robot are perceived correctly only if the trajectory is in conformity with that produced by an active movement. This law affects also the eye movements: the trajectory and the performance of visual tracking of a luminous moving point differ if the movement of this the point does not respect the law [12]. According to Viviani [13], the phenomenon finds its origin in the motor theory of perception: our perception of the movement is determined by our way to move and to act.

The power 2/3 law is considered as being a fundamental constraint of the CNS on the formation of the trajectories of the end-point of the gesture (e.g. the hand), in particular when performing rhythmic movements. This law also appears for more complex movements, concerning the whole body, as in locomotion [14, 15]. Confronted to a corpus of convergent experimental data, the power 2/3 law is regarded as an invariant of the trajectory of biological movements, impossible to circumvent and is often used as a criterion of evaluation of animated models [16, 17].

2.3 Local or Global Process: Maggie Shiffrar

The human body may be considered as a chain of rigid, articulated elements, giving to the body a non-rigid aspect. Shiffrar [18] attempted to determine how we can perceive a body as moving, and in particular how the visual system can integrate multiple segmental information on the moving body in order to perceive this movement as a single and continuous event. The assumption is that movement (animation) allows the establishment of a bond (rigid or not) between the various points. Beyond the perception of biological motion, this concept can be generalized to the perception of any

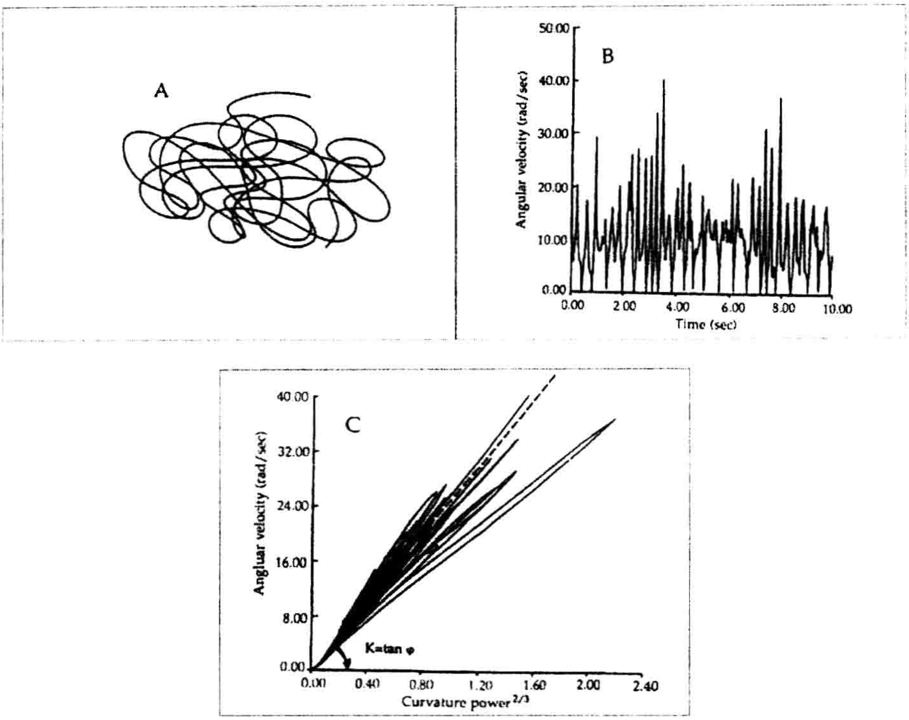

Fig. 2. Neither the trajectory of a graphic gesture (A) nor the time-course of angular velocity (B) reveal an invariant in the production of the gesture. On the other hand, the angular velocity is systematically proportional to the curvature of the trajectory raised at power 2/3 (C).

physical object and its identification by the visual system [19]. One of the questions tackled by the studies on perception of biological motion relates to the primary level of analysis, global or local (does each point count, or is it rather the general pattern which is relevant?). Shiffrar pleads for a global analysis: a group of dots representing a walker is not anymore recognized as such if it is reversed from top to bottom [20]; on the other hand, the walker is identified even if the characteristic points are drowned in a background of random dots [21]. It should finally be considered that more probably neither the shape (the space distribution) of the group of dots, nor the characteristics of the movement of the points are enough alone to define or categorize human movement. The capability of the visual system to extract a human form from motion is based on the space-time integration of the indices of BOTH form AND movement and must thus be regarded as a phenomenon of both local and global nature [18].

2.4 From Perception to Memory of Biological Motion: Perceptive Anticipation

This spatio-temporal integration does not only affect perception but also memory. Under certain conditions, our memory of the final position of a moving target which is abruptly stopped is distorted in the direction of the represented movement [22]. In the same way, our memory of a static view of a moving object or character is biased in the direction of the movement [23]. This phenomenon, called "representational momentum" attests of the "interiorisation" of the physical principle of inertia. It has been recently shown that the perception of complex biological movements is also affected by this phenomenon. Thus, even in case of disruption of a visual stimulation corresponding to a complex gesture, the perception of the event and its dynamics remain, resulting in a bias of the memory of the final posture, shifted in the direction of the movement [24].

The identification of actions and human body postures is a major task of our perceptive system, which depends on the point of view adopted by the observer. Indeed, visual recognition of a furtive human posture (presented during 60 ms) is facilitated by the previous presentation of identical postures, but only if these previous presentations are from close points of view [25]. In the same way, it is easier to assess the biological realism of a posture if the movement preceding this posture is also presented to the observer. Thus, one can anticipate the postures to result from a subject's movement, facilitating the identification of these postures [26]. Other studies on perceptive anticipation abound in this direction: it is enough for an observer to perceive the beginning of a gesture (i.e. a writing sequence) to correctly predict the nature of the incoming movements, and even if the produced series of letters do not form words [2].

2.5 Neural Substrates

Perception of biological motion is not a human exclusive capability: animals, and in particular monkeys or pigeons are also sensitive. This allowed, through a number of electrophysiological studies in primates (cell recording), to determine the concerned cortical zones. Oram and Perrett [28] showed that neurons of the temporal superior area respond to this kind of specific stimuli. Newsome and Paré [29] showed that monkeys can detect the direction of a movement when the level of coherence is as low as 1% or 2%. Destruction of a specific cortical zone (MT) increased this threshold of coherence to 10% or 20%.

Brain functional imaging now makes it possible to identify, on humans, zones of the brain specifically activated during perception and recognition of biological motion, and especially to determine the networks involved. The supero-temporal sulcus seems to play a particular role in the perception of biological motion [30] (Fig. 3). These studies made it possible to highlight the implication of a large number of zones, including of course those directly concerned with vision (in particular the lingual gyrus [31]); and the superior occipital gyrus [32] but also perceptive processes in general (temporal and parietal cortices), as well as other parts of the brain generally concerned with the control of movement (premotor cortex, lateral cerebellum), thus confirming the close link between the generation of movement and its perception [33].

Thus, for example, experiments of neuroimagery carried out in monkeys [34] and on humans [35, 36] showed that neurons known as "mirror neurons", located in the lower part of the premotor cortex are active both when one observes somebody

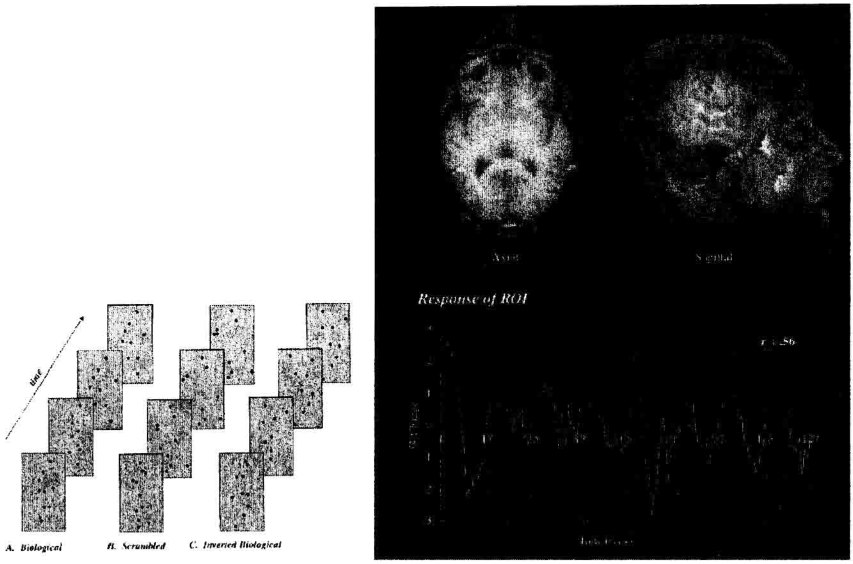

Fig. 3. The activity in specific zones of the brain (here the superior temporal sulcus) is modulated according to the presentation of a biologically compatible stimulus vs. a non compatible (scrambled or inverted) [30]

performing a gesture or when the subject is performing himself the gesture. This network would have a fundamental role in the processes of training by imitation and social communication [37]. These studies and those concerning the role of the ventral and dorsal flow of visual information in motion perception are clearly related.

3 Does Virtual Reality Have to Take into Account the Laws of Biological Motion?

At this point, I would like to emphasise the need, for people involved in VR, to know the main theories of perception and movement control, proposed throughout the 20th century by the psychologists and the physiologists. These theories are in three categories:

- Cognitive theories: they consider perception as a particular form of inferential process, although mainly unconscious. This point of view finds its origin in Helmoltz' work.
- Ecological theories, assuming that we (actors) have a privileged relation with the environment, and that we are directly in interaction with this latter, as proposed by Gibson [38].
- Computational theories, assuming that perception is primarily the result of data processing. Marr is one of the pioneers of this approach.