



INTELLI 2022

The Eleventh International Conference on Intelligent Systems and Applications

ISBN: 978-1-61208-977-5

May 22nd –26th, 2022

Venice, Italy

INTELLI 2022 Editors

Gil Gonçalves, FEUP, Portugal

Carsten Behn, Schmalkalden University of Applied Sciences, Germany

INTELLI 2022

Forward

The Eleventh International Conference on Intelligent Systems and Applications (INTELLI 2022) continued a series of events on advances towards fundamental, as well as practical and experimental aspects of intelligent and applications.

The information surrounding us is not only overwhelming, but also subject to limitations of systems and applications, including specialized devices. The diversity of systems and the spectrum of situations make it almost impossible for an end-user to handle the complexity of the challenges. Embedding intelligence in systems and applications seems to be a reasonable way to move some complex tasks from user duty. However, this approach requires fundamental changes in designing the systems and applications, in designing their interfaces and requires using specific cognitive and collaborative mechanisms. Intelligence become a key paradigm and its specific use takes various forms according to the technology or the domain a system or an application belongs to.

We take here the opportunity to warmly thank all the members of the INTELLI 2022 technical program committee, as well as all the reviewers. The creation of such a high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to INTELLI 2022. We truly believe that, thanks to all these efforts, the final conference program consisted of top-quality contributions. We also thank the members of the INTELLI 2022 organizing committee for their help in handling the logistics of this event.

INTELLI 2022 Chairs

INTELLI 2022 Steering Committee

Carsten Behn, Schmalkalden University of Applied Sciences, Germany

Stefano Berretti, University of Firenze, Italy

Leo van Moergestel, HU University of Applied Sciences Utrecht, Netherlands

Gil Gonçalves, University of Porto, Portugal

Marcin Paprzycki, Systems Research Institute, Polish Academy of Sciences – Warszawa, Poland

INTELLI 2022 Publicity Chairs

Mar Parra, Universitat Politècnica de València (UPV), Spain

Hannah Russell, Universitat Politècnica de València (UPV), Spain

INTELLI 2022 Committee

INTELLI 2022 Steering Committee

Carsten Behn, Schmalkalden University of Applied Sciences, Germany
Stefano Berretti, University of Firenze, Italy
Leo van Moergestel, HU University of Applied Sciences Utrecht, Netherlands
Gil Gonçalves, University of Porto, Portugal
Marcin Paprzycki, Systems Research Institute, Polish Academy of Sciences – Warszawa, Poland

INTELLI 2022 Publicity Chairs

Mar Parra, Universitat Politècnica de València (UPV), Spain
Hannah Russell, Universitat Politècnica de València (UPV), Spain

INTELLI 2022 Technical Program Committee

Azizi Ab Aziz, Universiti Utara Malaysia, Malaysia
Lounis Adouane, Université de Technologie de Compiègne, France
Ari Aharari, SOJO University, Japan
Bilal Ahmad, University of Warwick, UK
Zaher Al Aghbari, University of Sharjah, UAE
Antonios Alexos, University of California Irvine, USA
Sarrah Alqahtani, Wake Forest University, USA
Rachid Anane, Coventry University, UK
Zahra Ebadi Ansaroudi, University of Salerno, Italy
Olugbenga Moses Anubi, Florida State University, USA
Benjamin Aziz, University of Portsmouth, UK
Arvind Bansal, Kent State University, USA
Suzanne Barber, The University of Texas at Austin, USA
Carmelo Bastos-Filho, University of Pernambuco, Brazil
Rafael Batres, Tecnológico de Monterrey, Mexico
Carsten Behn, Schmalkalden University of Applied Sciences, Germany
Fayçal Bensaali, Qatar University, Qatar
Lyes Benyoucef, Aix-Marseille University, France
Giuseppe Berio, IRISA | Université de Bretagne Sud, France
Stefano Berretti, University of Firenze, Italy
Jonathan Berrisch, University of Duisburg-Essen, Germany
Mahdis Bisheban, National Research Council Canada (NRC), Canada
José Miguel Blanco, Masaryk University, Brno, Czech Republic
Francisco Bonin Font, University of the Balearic Islands, Spain
Lucas Botoni de Souza, Federal University of Technology - Paraná, Brazil
Frederic Bousefsaf, LCOMS | Université de Lorraine, France
Lars Braubach, Hochschule Bremen, Germany
Simeon C. Calvert, Delft University of Technology, Netherlands
Valérie Camps, Paul Sabatier University | IRIT, Toulouse, France

Carlos Carrascosa, Universitat Politècnica de València, Spain
Cesar Castelo-Fernandez, Institute of Computing | University of Campinas, Brazil
Dario Cazzato, Centre for Security, Reliability and Trust | University of Luxembourg, Luxembourg
Martin Cech, University of West Bohemia, Pilsen, Czech Republic
Chin-Chen Chang, Feng Chia University, Taiwan
Tongwen Chen, University of Alberta, Canada
Guilherme Conde, Federal University of Western Pará, Brazil
Angelo Croatti, University of Bologna, Italy
Daniela D'Auria, Free University of Bozen-Bolzano, Italy
Mohammed Dahane, Université de Lorraine, France
Chuangyin Dang, City University of Hong Kong, Hong Kong
Andrea D'Ariano, Roma Tre University, Italy
Jos De Brabanter, KU Leuven ESAT-STADIUS, Belgium
Toon De Pessemier, imec - WAVES - Ghent University, Belgium
Angel P. del Pobil, Jaume I University, Spain
Jens Dörpinghaus, Federal Institute for Vocational Education and Training (BIBB) / German Center for Neurodegenerative Diseases (DZNE), Germany
Nelson Duarte, CIICESI | ESTG | Politécnico do Porto, Portugal / IRIEM, Hong Kong
Arianna D'Ulizia, National Research Council - IRPPS, Italy
Chuck Easttom, Georgetown University, USA
Ahmed Ewais, Arab American University, Jenin, Palestine / Vrije Universiteit Brussel, Belgium
Tullio Facchinetti, University of Pavia, Italy
Ana Fernández Vilas, School of Telecommunication Engineering | University of Vigo, Spain
Stefka Fidanova, ICT-BAS, Sofia, Bulgaria
Manuel Filipe Santos, University of Minho, Portugal
Pedro Freitas, Universidade Católica Portuguesa, Portugal
Edgar Giovanni Cuzco Silva, Universidad Nacional de Chimborazo, Ecuador
Todorka Glushkova, Plovdiv University "Paisii Hilendarski", Bulgaria
Gil Gonçalves, University of Porto, Portugal
Sérgio Gorender, Federal University of Bahia, Brazil
Mohammad Goudarzi, University of Melbourne, Australia
Javier Gozalvez, Universidad Miguel Hernandez de Elche, Spain
Emmanuelle Grislin-Le Strugeon, LAMIH | Université Polytechnique Hauts-de-France (UPHF), France
Ibrahim A. Hameed, Norwegian University of Science and Technology (NTNU), Norway
Ridewaan Hanslo, University of Pretoria, South Africa
Wladyslaw Homenda, Warsaw University of Technology, Poland
Tzung-Pei Hong, National University of Kaohsiung, Taiwan
Wei-Chiang Hong, School of Education Intelligent Technology - Jiangsu Normal University, China
Matin Hosseini, University of Louisiana at Lafayette, USA
Christopher-Eyk Hrabia, Technische Universität Berlin | DAI-Labor, Germany
Chih-Cheng Hung, Kennesaw State University - Marietta Campus, USA
Syed Muhammad Zeeshan Iqbal, BrightWare LLC, Riyadh, Saudi Arabia
Zahid Iqbal, University of Porto, Portugal
Ajune Wanis Ismail, Universiti Teknologi Malaysia, Malaysia
Raheleh Jafari, School of Design | University of Leeds, UK
Anubhav Jain, Telstra, India
Juergen Jasperneite, Fraunhofer IOSB-INA, Germany
Thomas Jell, Siemens Mobility GmbH, Germany

Andrés Jiménez Ramírez, University of Seville, Spain
Maria João Ferreira, Universidade Portucalense, Portugal
Mihaela Juganaru, IMT - Mines de Saint Etienne, France
Janusz Kacprzyk, Systems Research Institute - Polish Academy of Sciences, Poland
Ryotaro Kamimura, Tokai University, Japan
Keiichi Kaneko, Tokyo University of Agriculture and Technology, Japan
Stelios Kapetanakis, University of Brighton, UK
Mehdi Kargar, Ted Rogers School of Management | Ryerson University, Canada
Alexey Kashevnik, SPIIRAS, Russia
Okba Kazar, University of Biskra, Algeria
Alireza Khanteymooori, Universitätsklinikum Freiburg, Germany
Leoneed Kirilov, Institute of Information and Communication Technologies - Bulgarian Academy of Sciences, Sofia, Bulgaria
Sotiris Kotsiantis, University of Patras, Greece
Boris Kovalerchuk, Central Washington University, USA
Akmaral Kuvatbayeva, International IT University, Almaty, Kazakhstan
Tobias Küster, DAI-Labor / Technical University of Berlin, Germany
Victoria Lapuerta, Universidad Politécnica de Madrid, Spain
Antonio LaTorre, Universidad Politécnica de Madrid, Spain
Frédéric Le Mouël, Univ. Lyon / INSA Lyon, France
Deok-Jin Lee, Kunsan National University, South Korea
George Lekeas, City Universty - London, UK
Maurizio Leotta, University of Genova, Italy
Chanjuan Liu, Dalian University of Technology, China
Francesco Longo, University of Calabria, Italy
Daniela López De Luise, CI2S Labs, Argentina
Majdi Maabreh, The Hashemite University, Jordan
Francesca Maridina Mallocci, University of Cagliari, Italy
Harald Mayer, JOANNEUM RESEARCH Forschungsgesellschaft mbH, Austria
René Meier, Hochschule Luzern, Germany
António Meireles, GECAD - Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development, Portugal
Jérôme Mendes, Institute of Systems and Robotics (ISR-UC), Portugal
Márcio Mendonça, Federal University of Technology - Paraná (UTFPR), Brazil
Tarek Menouer, Umanis Research & Innovation, France
Jair Minoro Abe, Paulista University & Institute of Advanced Studies | University of São Paulo, Brazil
Dusmanta Kumar Mohanta, Birla Institute of Technology, India
Jose M. Molina, Universidad Carlos III de Madrid, Spain
Vítor Monteiro, University of Minho, Portugal
Ceci Morales, iRobot, USA
Fernando Moreira, Universidade Portucalense Infante D. Henrique, Portugal
Paulo Moura Oliveira, UTAD University, Vila Real / INESC-TEC- Technology and Science, Porto, Portugal
Debajyoti Mukhopadhyay, Mumbai University, India
Muddasar Naeem, ICAR-CNR, Naples, Italy
Filippo Neri, University of Naples, Italy
Pranav Ajeet Nerurkar, NMIMS University, Mumbai, India
Dinh-Luan Nguyen, Michigan State University, USA

Thanh-Tuan Nguyen, HCMC University of Technology and Education, HCM City, Vietnam / University of Toulon, CNRS, LIS, Toulon, France
Alex Norta, Tallinn University of Technology, Estonia
Cyrus F. Nourani, akdmkrd.tripod.com, USA
Kenneth Nwizege, Ken Saro-Wiwa Polytechnic, Nigeria
Jin Dong, Oak Ridge National Laboratory, USA
Michel Occello, Université Grenoble Alpes, France
Krzysztof Okarma, West Pomeranian University of Technology in Szczecin, Poland
Ana Oliveira Alves, Coimbra Polytechnic - ISEC & Centre of Informatics and Systems of the University of Coimbra - CISUC, Portugal
Joanna Isabelle Olszewska, University of West Scotland, UK
Yash-Vardhan Pant, University of California, Berkeley, USA
Marcin Paprzycki, Systems Research Institute / Polish Academy of Sciences - Warsaw, Poland
Carla Pereira, School of Technology and Management / INESC TEC, Portugal
Isidoros Perikos, University of Patras, Greece
Goharik Petrosyan, International Scientific-Educational Center of the National Academy of Sciences, Yerevan, Armenia
Agostino Poggi, Università degli Studi di Parma, Italy
Marco Polignano, University of Bari "Aldo Moro", Italy
Filipe Portela, University of Minho, Portugal
Catia Prandi, University of Bologna, Italy
Dilip Kumar Pratihar, Indian Institute of Technology Kharagpur, India
Radu-Emil Precup, Politehnica University of Timisoara, Romania
Shahnawaz Qureshi, National University of Computer and Emerging Sciences, Pakistan
Ahmed Rafea, American University in Cairo, Egypt
Giuliana Ramella, National Research Council (CNR) - Institute for the Applications of Calculus "M. Picone" (IAC), Italy
Aurora Ramírez, University of Córdoba, Spain
Chakroun Rania, National School of Engineering of Sfax | Advanced Technologies for Image and Signal Processing (ATISP) Research Unit, Sfax, Tunisia
Radha Reddy, CISTER Research Center | ISEP | FEUP, Porto, Portugal
Carlos Renato Vázquez, Tecnológico de Monterrey, Mexico
Rabiâ RIAD, Ibn Zohr University, Morocco
Fátima Rodrigues, Institute of Engineering | - Polytechnic of Porto, Portugal
Daniel Rodriguez, University of Alcalá, Spain
Federica Rollo, University of Modena and Reggio Emilia, Italy
Amirreza Rouhi, Politecnico di Milano, Italy
Shah Rukh Humayoun, San Francisco State University, USA
Alexander Ryjov, Lomonosov Moscow State University | Russian Presidential Academy of National Economy and Public Administration, Russia
Fariba Sadri, Imperial College London, UK
Mohammad Saeid Mahdavinejad, Kansas State University, USA
George E. Sakr, St Joseph University of Beirut, Lebanon
Bilal Abu Salih, Curtin University, Australia
Demetrios Sampson, Curtin University, Australia
Christophe Sauvey, LGIPM | Université de Lorraine, France
Alessandra Scotto di Freca, Università di Cassino e del Lazio Meridionale, Italy
Chantal Soulé-Dupuy, University of Toulouse Capitole, France

Francisco Souza, Radboud University, the Netherlands
Sashank Sridhar, College of Engineering Guindy - Anna University, India
Nick Taylor, Heriot-Watt University, UK
Mark Terwilliger, University of North Alabama, USA
Supphachai Thaicharoen, Srinakharinwirot University, Bangkok, Thailand
Pei-Wei Tsai, Swinburne University of Technology, Australia
Alexandros Tzanetos, Université de Sherbrooke, Canada
Berna Ulutas, Eskisehir Osmangazi University, Turkey
Paulo Urbano, Universidade de Lisboa - BioISI, Portugal
Leo van Moergestel, HU University of Applied Sciences Utrecht, Netherlands
Jan Vascak, Technical University of Kosice, Slovakia
Costas Vassilakis, University of the Peloponnese, Greece
Anna-Maria Velentza, University of Macedonia, Thessaloniki, Greece
Chong Wang, S&P Global, USA
Minjuan Wang, San Diego State University, USA
Yifei Wang, Georgia Institute of Technology, USA
Kanoksak Wattanachote, Guangdong University of Foreign Study, China
Dietmar Winkler, TU Wien | CDL-SQI, Vienna, Austria
Stefanie Wuschitz, Miss Baltazar's Laboratory, Vienna, Austria
Mudasser F. Wyne, National University, USA
Maria Gabriella Xibilia, University of Messina, Italy
Peng Xu, Technical University of Munich (TUM), Germany
Longzhi Yang, Northumbria University, UK
Leila Zemmouchi-Ghomari, Ecole Nationale Supérieure de Technologie, ENST, Algiers, Algeria

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission to reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

| | |
|---|----|
| The Visual Consciousness Space: A Mathematical Topological Proof of the Irreducibility of Consciousness to Physical Data <i>Igor Reznikoff</i> | 1 |
| Smart Factory Automation for Robotic Production of Satellite Formations and Constellations <i>Klaus Schilling</i> | 7 |
| A Deep Learning based Unoccupied Parking Space Detection Method for City Lots <i>Hamid Reza Tohidypour, Yixiao Wang, Panos Nasiopoulos, and Mahsa T. Pourazad</i> | 11 |
| Semantic Patterns to Structure TimeFrames in Text <i>Nour Matta, Nada Matta, Nicolas Declercq, and Agata Marcante</i> | 16 |
| Linear Fuzzy Space Based Framework for Air Quality Assessment <i>Endre Pap, Dorde Obradovic, Zora Konjovic, and Ivan Radosavljevic</i> | 24 |
| Reconfigurable Digital Twins for an Industrial Internet of Things Platform <i>Eliseu Pereira, Maria Arieiro, and Gil Goncalves</i> | 30 |
| Applying Deep Learning Techniques in Automated Analysis of Echocardiograms, CMRs and Phonocardiograms for the Detection and Localization of Cardiac Diseases <i>Arvind Bansal and Racheal Mukisa</i> | 36 |
| Steps towards the Modeling of Animal Vibrissa Modes Using Adaptive Control <i>Carsten Behn, Moritz Scharff, and Lukas Merker</i> | 43 |

The Visual Consciousness Space:

A Mathematical Topological Proof of the Irreducibility of Consciousness to Physical Data.

Iegor Reznikoff

Emeritus Professor, Département de Philosophie
Université de Paris Ouest Nanterre
Paris, France
e-mail: dominiqueleconte@yahoo.fr

Abstract—We show here that what we call ‘visual space of consciousness’, the space of what we see, is a specific space different from the purely physical one and that its properties imply that it cannot be reduced to or deduced from physical laws. Some biological points are also briefly considered. The arguments are of logical, mathematical and physical character, and although elementary, they require a careful reading. There is no need to define consciousness; we only observe some of its properties, namely geometric and topological properties of visual consciousness, and show that these properties cannot be based on physics only. Now, if a part of consciousness cannot be grounded on physics only, it is the same for consciousness as a whole and we speak of the irreducibility of consciousness to physical data. We do not consider philosophical questions or issues; in a simple physical and mathematical frame, we give a logical proof of this irreducibility. Elements for a formal mathematical, logical proof are mentioned at the end of the paper. Concerning intelligent systems, this paper is rather a challenge: it reveals a limit for wished achievements of intelligent systems. Moreover, the idea of non reducibility of consciousness to physical properties may help conceive foundations for studies in possible contact-less brain-computer interface.

Keywords—consciousness; vision; physics; reality; continuity; logic.

I. INTRODUCTION

The main purpose of this work is to give a proof of the non reducibility of consciousness to physical data. In order to treat the problem precisely and have clear definitions, we essentially limit the question to the *visual space*, i.e., to the space we see (when looking at something), so that we need not define *consciousness*. If a part – the visual one – of consciousness cannot be founded on physics only, it is the same for consciousness as a whole and we speak of the irreducibility of consciousness to physical data. In Sections II and III, notions of *visual space* and *irreducibility* are respectively defined; then we study two main properties of the visual space, namely its *continuity* (Section IV) and *unity* (Section VI). Since regarding unity the biological level is concerned, this point is briefly discussed (Section VII). The last section gives elements for a formal mathematical and logical proof (which is out of the scope of this paper). Finally, we conclude with a short historical perspective. The strictly scientific approach of this work and difficulties we met may give new ways to study the possibility (or not) of

contact-less brain and computer interface, which is certainly a tricky subject.

A first shorter version of this paper appeared in a hardly accessible Journal [1] (see also [2]).

II. THE CONSCIOUSNESS SPACE

For a given observer, let A be the space of ‘physical reality’ as known by physics, the ‘real’ space of matter with what is included in it: moving atoms, particles and waves; and let B be the observer’s brain regarded as a space, with its physiological and neuronal activity (of course $B \subset A$); then let C be the space of the observer’s perceptive consciousness: what he sees, hears, touches, etc. considered as a space. There are, of course, further levels of consciousness, in particular a witness consciousness: the one that sees, hears, etc. (not to speak of the thinking one). But, here, we consider only perceptive consciousness, what is seen (heard, etc.), and, more precisely, the visual space in the case of vision (resp. the spaces of what one hears, touches, etc. for the other senses). For simplicity and because of its obvious geometric appearance, we confine our remarks mostly to the visual space, but the same points can be made about other spaces of perception (For a very interesting approach of consciousness of sound, see [3] and [4]). In what follows, we will speak of A , B , and C also as being respectively the ‘*real*’ or *physical space*, the brain space and the *consciousness space* with its visual sub-space.

Between a part of $A - B$ (elements of A that are not in B) and B there is a map, say f , which to physical events in this part of $A - B$, through the perceptive channels, associates reactions in the brain space B . For instance, a photon flux received by the eyes creates an activity in the optical nerve and then in the brain. To this activity at the level B corresponds in general a representation in C ; let g be this correspondence between brain activity and its representation (image) in the consciousness space. There is, therefore, a correspondence from A to C defined by $h = g \circ f$. We have the following diagram (Figure 1).

If we limit C to perceptive consciousness – as we are doing here – we could expect to have $g \circ f(A - B) = C$, but this is not the case since we are going to show that there are properties of C which do not proceed from properties of A . Note that f is not injective, since different points in A cannot always be distinguished in B , and even less in C . One might discuss, of course, the precise domains of definition of f and g – since there are no, for all events in A

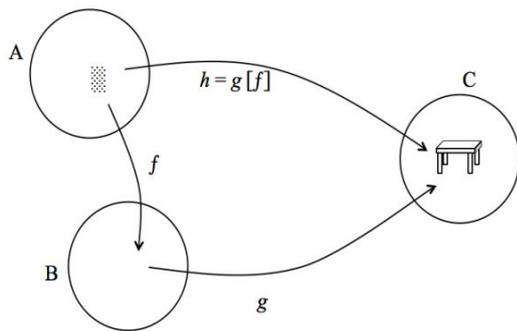


Figure 1. Spaces A, B, C as presented in the text.

or in B, corresponding reactions or representations in B or C, respectively, but this is of no importance here. Also we did not give a precise mathematical definition of the maps since it is not even clear what the elements in A, B and of course in C are; one should probably rather speak of a map between some subsets of A and subsets of C, or even better in the language of categories; for an approach of some properties of consciousness in terms of categories, see [5] (I am thankful to the mathematician G. Choquet who mentioned this remarkable work). For sake of simplicity, we keep the elementary formulation in the language of set theory, also because we shall have a logical approach referring to some axioms of set theory. The given correspondences f and g are however clear enough to say that h gives a representation or an image of the reality A in the consciousness space C. For instance: to a subset T of $A - B$, a subset of particles reflecting a flux of photons, is associated, through the brain space B, a representation $h(T)$ of T, in C, say a table; this table is a representation of external physical data (the set of all particles, waves, etc. concentrated in the given space we call *table*). We are going to study some properties of the space of such representations, what we call the consciousness space C.

Proposition 1. The space C is a specific space in itself, different from A and B.

Proof. We have $C \neq A$, since the representation of $A - B$ in C depends on B (and g), whereas $A - B$ does not. For example if B is disturbed or injured, clearly so will be C (for instance a person sees two tables or none), while $A - B$ has not changed (there is still one table). But also $C \neq B$. What we see is not the brain, nor the activity of the brain. The same arguments show, moreover, that $C \cap A = \emptyset$ and $C \cap B = \emptyset$. Stated otherwise, the proposition says that

the consciousness space is different from the real physical space and from the brain space.

Given an object X in A, we do not see X in itself but only the result in C of a physiological and neuronal activity in B created by photons coming from X. This is well known, but the unconscious identification of $h(X)$ with X, of what we see with the material physical world, is so strong and so widespread – even for those who have read Kant – that it is necessary to restate clearly: there is a proper space

of consciousness and the picture that one sees is not the physical world. It is merely a representation of a set of particles and waves in the consciousness space; it does not mean, however, that this space is itself a ‘set of particles’: this is the point, the visual space is a proper non discrete space (see below).

An interesting question is that of the dimensionality of the consciousness space. For the visual space, one usually counts three spatial dimensions and three dimensions for colours. But we do not see a mixture of three colours; we see at least seven colours and their varieties and mixtures (For a topological approach of visual and colour spaces, see [6]-[8]). And what about the dimensions of the spaces of what we hear [4], smell or touch? Here, however we do not consider this question.

III. ON IRREDUCIBLE PROPERTIES OF THE CONSCIOUSNESS SPACE

It is clear that the consciousness space has properties that are not as such in the physical space: this is the case, for instance, for colours, which are indefinable without the direct experience of seeing them (a person blind from birth can have no idea of what *green* means, though he might associate other feelings with this word). However, although the qualities (*qualias*, see [9]) of a colour cannot be given physically, they do correspond to physical data: a green object reflects the light waves with a frequency that we see as ‘green’. We say that this property of being green can be *reduced* to physical properties.

The question then arises whether there are properties of the consciousness space that cannot be reduced to real physical ones. If so, we speak of *irreducible* properties.

IV. CONTINUITY OF THE CONSCIOUSNESS SPACE

A remarkable property of the visual space (but it is true also of other perceptions, although the matter is more difficult to formulate) is its continuity. We understand continuity in the mathematical sense, but the following elementary definition is sufficient here: in the visual space, there are no gaps or moving separate points; e.g., an ordinary white sheet of paper appears uniformly, permanently white and still, for at least a while. In contrast, physical reality at the atomic level is essentially discrete, non uniform, never motionless, and full of collisions; it does not mean that there are holes of energy or whatever, but we number atoms, electrons and various particles. With modern laser and other technologies, an isolated electron can be observed [10][11], and with the Nitrogen Vacancy (NV) nanotechnology it is possible to produce sources of isolated photons [12]. Moreover, particles and waves are in perpetual movement. Of course, at our macroscopic level, we can use a magnifying glass and discover other aspects of what our eyes and consciousness did not see before, but the image remains continuous; there are no holes in the space of visual consciousness. Thus, we have the question: how can a discrete moving atomic reality be represented in a continuous way (both in space and time)? The usual explanations about such questions concern the macroscopic

level of perception. A typical example is given by the continuous appearance of a discrete pile of stones seen from a distance: it seems to be a white continuous spot. The usual argument is that the discreteness is too subtle to be perceived. The question then arises: to be perceived by what? And where? Indeed this gives no explanation at the atomic level, since the perceptions are transmitted and received (if we remain in a purely physical world) by discrete processes of particles and moving waves, particularly photons, and charges. How does this produce a continuous image, and where does this image appear? Certainly not in a physical space of particles, nor in the neuronal brain; the neurones transmit physical information up to consciousness which produces the continuous image we see; this is also a proof that C , in no way belongs to the physical or brain space, and is a specific space of non-material, non-physical nature. Something to be perceived needs a perceiver and here the perceiver's visual continuous space cannot be reduced to physics because of the argument above. If everything were created, transmitted and received by physical spaces, it would remain permanently moving and discrete. As we say in French, "the most beautiful girl in the world cannot give more than what she has": physics gives no more than physics.

Proposition 2. The property of continuity of the consciousness space C is not reducible to or even explainable in terms of physical reality.

Against this statement, there is also the argument that if there were holes in the visual space we simply would not see them, since we obviously cannot see what we do not see (!). The discrete structure, therefore, cannot be perceived. However, this argument again supposes that something is perceived and already presupposes a perceiver: who is the *we* in the sentence above and who or what is seeing? In such answers, it is assumed that something (somebody) already sees or does not see; and the question remains how, in the final analysis, a 'continuous' space of vision can exist and where it can be located. Since it cannot be based on physics only, the conclusion is straightforward: continuity is a creation of consciousness. And, here, we come to a purely mathematical and logical consideration: continuity is not definable from discreteness and finite considerations, and cannot exist in a finite numbered domain. But physical reality – in a bounded domain at least – is finite. The property of continuity is, therefore, indeed irreducible to any physical reality, unless the notion of continuous field be introduced, which is a very theoretical and problematic notion that we discuss below.

A related important remark is about the relevance of the so called Nyquist-Kotelnikov-Shannon Theorem, to this problem of continuity. Briefly speaking, this theorem shows that it is possible to restore a regular continuous wave by knowing a finite (but sufficient) number of points on this wave (see [13] and [14]). Strictly mathematically, this result requires Fourier transforms, but even at a simpler mathematical level, our *perceptual* problem is not concerned: it is not because $9 \times 40 = 360$ that we see 360

points in a set of 9 rows of 40 points; we need to count or calculate. Thought is needed, perception is not sufficient; while here, we are concerned by visual perception only.

A. Commentary

One can discuss whether the property of continuity is needed to characterize the visual space; for instance, is the property of *density* not sufficient (as for the line of rational numbers)? Let us recall that density means that between two points there is always a third one. The above irreducibility argument remains in force even if we assume density; since density implies infinity, even in an interval or bounded space; indeed, the fact that between two points there is a third one implies that between these two there is an infinity of points. An absolute proof of the continuity of the visual space in the strong mathematical meaning is certainly not possible because it requires high technical considerations of infinite character – let us recall that the continuity of a space implies that its infinity is not countable, which means that it is bigger than the infinity of the set of integers (it is said to have the *power of the continuum*). Such considerations are purely theoretical and certainly beyond any experience. But there is another strong epistemological argument for attributing continuity to the visual space. This argument comes from answering the question: how did the concepts of geometrical (Euclidian) space and precisely of continuity appear? How the geometrical line was and is understood to be continuous?

The notions of *geometrical space* and line appear, of course, in and from our visual space and visual experience (connected with that of movement and touch for three-dimensional awareness). Moreover, all our intuition of space geometry in the plane comes essentially from our visual space which, as we know, until the discoveries of Relativity Theory, was considered in its Euclidian formulation to be absolute (from the physical point of view at least). The notion of *continuity* proceeds as well from our visual experience, the best notion of a continuous line or surface being probably given by a surface of water: there are no holes or separations. That a segment of a straight line has infinitely many points (because it is dense) is readily understood and has been understood since Antiquity as well as the (intuitive) continuity of the line. And it is most remarkable that children, from visual experience, easily understand the notion of a (straight) line as well as its potential infinity and its continuity as being with 'points everywhere' so that there are no holes left. Of course the notions of closeness, or of going through are also related to our experience of movement and touch, but, finally it is by reasoning on the geometric line, which belongs to and comes from our visual space (so that it can be drawn), that the theory of this geometric line has been worked out. Also, let us note that in our visual space all (necessary macroscopic) movements are continuous: it is impossible to join two points without passing somewhere in-between, while this is not the case at the atomic quanta level. It is, therefore, quite reasonable to consider that our intuition and understanding of the visual space *demonstrate it to be a continuous space*.

What we said about spatial continuity can be repeated concerning the continuity of the visual space in time (and, more generally, of perception in time). The visual space lasts in a continuous way as does a continuous movement; while at the atomic level, in duration, there is no continuity at all. However, our perception of time is continuous and has led to a theoretical treatment of time which identifies it with the geometric line (for a study on time based on a distinction of physical and mind levels, see [15]). This continuity in time is closely related to what appears to be an even more remarkable property of consciousness and particularly of the visual space, namely its unity (see below, Section VI).

B. The consciousness space as a field

The only physical approach to continuity is given by the notion of field. For instance, an electromagnetic or gravitational field is assumed to be defined and active everywhere in the physical domain where it acts, and this *everywhere* is understood to be continuous since the space where the field is active is mathematically considered to be the three-dimensional space R^3 . This is a purely mathematical and theoretical formulation: we can only verify that the field acts on every particle or object appearing in the domain, and experience can show no more. But to assume continuity and R^3 allows us to use the mathematical infinitesimal calculus with all its tremendous power. However, we claim, after the discussion above, that this geometrical approach is a creation of consciousness and particularly of consciousness of the visual space, since there is no other evidence for such a geometrical and topological continuous conception. It is not and cannot be given by direct physical experience which is finite. For us, therefore, this geometry is not in A, but in C, and then induced from C to the theoretical, mathematical treatment by using the space R^3 containing a theoretical model of A. There is a brain electromagnetic field [16], but we do not see this field.

But, since this notion of continuous field actually exists in physics – be it created by and conceived from visual consciousness – we may say that consciousness is indeed a field. And just as a movement of electrons creates an electromagnetic field, we may conjecture that intense brain activity – of purely physical character at the atomic level of particles and waves – creates a field of consciousness: the greater the brain activity, the richer the field of consciousness. This field is, of course, not physical, since $A \cap C = \emptyset$ as we have seen, the space C being a specific one. Nor it is simply reducible to a known physical field, we certainly do not see an electromagnetic or gravitational field. Moreover, in visual consciousness, we can isolate forms, colours, objects, etc., while even if there are different wave lengths, etc. in physical fields, it is consciousness that extracts the mentioned forms, properties or elements from the visual consciousness space we see. There is nothing analogous or even expressible concerning physical fields. The property of seeing separate objects in the unity of the whole visual picture corresponds to the *Comprehension Axiom* of set theory: given a set E and a property P, the subset of elements of E verifying the property P, exists. It

cannot be deduced from physics. It seems to be a fundamental property of consciousness related to the *a priori* capacity of consciousness to pay attention. Moreover, if we consider the whole perceptive consciousness, there is no homogeneity between visual pictures, acoustic perception or touch.

If we consider the spaces of what we see, hear, smell, taste or touch, as different subspaces of the whole consciousness space, the non-homogeneity of these subspaces is a quite peculiar fact, bearing in mind that they are all produced by the same kind of neuronal activity, since there seems to be no difference between the neurones of different perception areas in the brain. How can the same kind of neuronal activity produce such different worlds of perception? This question could yield another proof of the irreducibility of the consciousness space to the physical one. Of course, the scales of various physical data producing the perceptions are quite different, e.g., the scales of light waves, molecules (for the smell) or sound waves, but this does not explain the complete non-homogeneity of the corresponding subspaces of consciousness, whereas they are held together in a remarkable unity: I smell the rose that I see. This is a specific property of consciousness.

But, since the emergence of space consciousness comes mostly from an intense brain activity of quantum electromagnetic nature, the relationship between such a field and the field of consciousness has to be investigated not only for isolated phenomena (for instance the fact that different frequencies of light produce different colours in the visual consciousness) but in the whole. Why would not a special intense physical activity – in the brain – create a field of different – non-physical – nature? Clearly, the intermediate biological level appears as an essential one.

V. THE OBJECTIVITY OF THE CONSCIOUSNESS SPACE C

A peculiarity of the consciousness space is that it can be studied essentially only from inside, by itself: only consciousness knows consciousness. And “if you want to know my consciousness, look in yours” sounds as a wise saying. Thus, we come to this important statement:

Proposition 3. The properties of C can be seen by everybody: its study is therefore perfectly objective.

Note that, here, the word 'objective' has the same meaning as in natural sciences, e.g., physics, since everything we know is known from our perceptive consciousness, and everything we look at – for instance the position of a needle in a measuring apparatus – is seen in our visual consciousness, that is in C. If two persons see the same object (the needle at a given position) it is because it is the ‘same’ object in their respective visual spaces. Although the meaning of the word *same* cannot be explained, this meaning is based on a universal understanding without which no communication would be possible. Here, appears the common but meaningless question whether we all really see the same colours or objects: is the red that I see the same as the one you see? The question is meaningless because it cannot be verified, but *the simpler the hypothesis, the better*

it is, and the simplest is to consider that indeed we have essentially the same consciousness. But, as mentioned above, philosophical discussions are not considered here (for a discussion, see [17]).

VI. UNITY OF THE VISUAL CONSCIOUSNESS SPACE

One of the most remarkable properties of consciousness space – and, moreover, difficult to understand – is its unity, that is the capacity that consciousness has to gather perceptions as a whole; from a multiplicity of independent nervous impulses and neuronal processes, consciousness produces a unified whole. We do not have consciousness of separated elements, but always of a coherent whole, even when looking at an isolated object.

This unity principle is the following: given separate elements x_1, \dots, x_n , it is the actual capacity to conceive their totality, i.e. the set $\{x_1, \dots, x_n\}$. It is remarkable that this corresponds to an axiom of set theory; logically, this property is not reducible. It cannot be deduced simply from the existence of x_1, \dots, x_n as separate elements. Therefore, it is not physically explicable, unless of course it is implicitly assumed (which is often the case, for instance when one assumes that things are somehow and somewhere ‘observed’ before any consciousness has been introduced). In particular, the argument that unity results simply from the simultaneity of neuronal processes in some centre of the brain is doubly inconsistent. First, because the notion of simultaneity is meaningless without the notion of *now* or the notion of *at the same time as*, which presupposes a reference and a clock and therefore an observer, i.e., a consciousness that grasps this simultaneity, this very notion introduces already an observer, it is not an absolute notion. And the second inconsistency is that simultaneity presupposes certainly the comparison of at least two elements and, hence, the notion of totality, be it only of the set $\{x_1, x_2\}$ as a whole. Therefore, to have the notion of simultaneity we already need that of unity; it is impossible to avoid circularity. The simultaneity of physical events is perhaps necessary for consciousness of unity, but not sufficient to explain it.

But, even if the notion of simultaneity is given, the probability that all the possible visible ‘dots’ of our visual neurology (e.g., retina) are grasped together in a coherent unity (their number can be estimated of the order of 10^7), this probability is of order of $2^{(10^7)}$ (2 to the power of 10 to the power 7), which is well beyond any physical meaning even at the level of light-wave length. Unity cannot emerge ‘by chance’; moreover, it is permanent, continuous in time. The probability for this continuing unity is physically without meaning.

This capacity of totalization, this gift of perceptive consciousness, is certainly one of its most important properties and unity may be the most characteristic property of consciousness. Consciousness unifies elements that otherwise are not related; from this comes what is called *meaning*. But we are not discussing this here any further.

As we have seen, the unity principle has, of course, no equivalent in physics; theoretically, it has to be borrowed

from logic. The property of unity, say of visual consciousness and of the space C , is irreducible.

Proposition 4. The unity of consciousness is not reducible to physical properties.

The question then arises how far logical arguments can be used in physics, biology and matters of consciousness. But, if one looks at a deductive science, rigorously founded, the logical and mathematical arguments are hitherto unavoidable. Of course, a science can be very rich as a descriptive one, but the claim is now: is it possible to ‘explain consciousness’ by neuronal and finally from purely physical processes? Since in our attempt, *explain* means *deduce* or *reduce*, the argument needs to avoid circularity, therefore a careful logical examination is needed, which we have attempted above: the unity of the visual space cannot be reduced to or explained by physics without circularity since a notion of unity is needed beforehand.

VII. BIOLOGICAL UNITY

If at some level the property of unity is needed and has to be introduced as such, then, it could be given already at a different level. It is natural to assume this unity, as we have seen, as one of the characteristic properties of consciousness, but it could be attributed already at the biological level. One often speaks of the ‘unity of the cell’. Is it not at this elementary biological level that a principle of irreducible unity has to appear?

That such unity is necessary as a global *principle* in biology is simply shown by the same argument as the one given for the impossibility of a random unity of the visual space. Suppose a biological organism of about 10^9 (of about 10 to the power of 9) components (e.g., molecules); the probability that all these components should behave *together* in the right way in order to constitute a biologically viable unity, this probability is at *least* of the order of $2^{(10^9)}$ (2 to the power of 10 to the power of 9), which is, as we have seen before, beyond any physical meaning even at the atomic level: the age of the universe would not be sufficient for even one cell to have a chance to exist, not to speak of a more complex organism.

However, even if a biological property should normally appear and be stated before properties of consciousness (and moreover could explain some of its aspects), we have a knowledge of the visual space, of its continuity and unity, certainly clearer, at least in its immediacy, than an as yet unformulated principle of unity in biology.

VIII. FORMAL APPROACH

For a formal, strictly deductive logical approach, we need different levels of axioms, laws and data, so that the following levels have to be distinguished.

1. The Logical level needed for mathematics. This introduces axioms, e.g., the *Axiom of Totality*: given x_1, \dots, x_n the set $\{x_1, \dots, x_n\}$ exists. The *Comprehension Axiom*: given a set E and a

property P, the subset of elements of E verifying the property P, exists. And finally an *Axiom of Infinity*.

2. The Mathematical level: theory of Real Numbers and Analysis. Logical axioms are intended for mathematical notions and reasoning.
3. The Physical level with its proper axioms and laws. The notions of continuity etc. are borrowed from the level 2.
4. The Biological level (this level is not really needed here).
5. The Consciousness level.

It is important to stress that we need not define consciousness (which would be a big challenge since consciousness is irreducible to other levels); we only observe some properties of visual consciousness, i.e., of what we see. But for continuity, unity and consciousness of seeing various objects, we need axioms analogous to the axioms above; these axioms are necessary to explain the mentioned properties of the visual space, and necessary for a deductive construction showing rigorously the irreducibility of consciousness to other levels. Since these axioms are not given by physics, clearly the level of consciousness is not at the physical level and cannot be deduced from physics.

IX. CONCLUSION

That consciousness space is relatively independent from external physical reality is a classical statement [18]. For Plato, Consciousness precedes Matter (as we learn from the *Timaeus*, 34c), the same for Indian classical religious philosophy; Kant's thoughts on this topic are well known (*Kritik der reinen Vernunft*), but it is worth quoting Berkeley: "*The proper objects of sight not without the mind; nor the images of anything without the mind*" and also "*Images in the eye are not pictures of external objects*" [19]. Here, we have simply shown that this relative independence of the consciousness space and its specific nature can be proved convincingly. The different points on which this proof is based, were studied in successive Sections of this paper; they give a serious challenge for intelligent machine approaches (see also [20]), but may suggest new rationale and reflection on such approaches.

REFERENCES

- [1] I. Reznikoff, "The consciousness space," Bio-Math, tome XXXX (n° 158), Paris, pp. 1-7, 2003.
- [2] I. Reznikoff, "A Logical and Topological Proof of the Irreducibility of Consciousness to Physical Data", <https://arxiv.org/abs/2110.14598> [retrieved: March,2022].
- [3] R. Casati and J. Dokic, *La Philosophie du Son (Philosophy of Sound)*, Nîmes (France), 1994.
- [4] I. Reznikoff, "On primitive elements of musical meaning," <http://www.musicandmeaning.net/issues/showArticle.php?artID=3.2>, 2005 [retrieved: March,2022].
- [5] A. Ehresmann, <https://vbm-ehr.pagesperso-orange.fr/> [retrieved: March,2022].
- [6] C. Pélissier, *Aspects physico-mathématiques de la vision de la couleur*, Thèse de Doctorat d'Etat, (Physical-mathematical aspects of color vision, Doctoral Thesis), Université Paris VI, 1982.
- [7] C. Pélissier, "Théorie unifiée de Young – Helmholtz – Hering, topologie induite, phénomènes d'adaptation, notion de temps et d'énergie" (Young – Helmholtz – Hering unified theory, induced topology, adaptation phenomena, notion of time and energy) in *Actes du 5e Congrès de l'Association Internationale de la Couleur, Mondial Couleur, Monte-Carlo, 1985*.
- [8] C. Pélissier, *Couleurs et Temps, de la Physique à la Phénoménologie (Colours and Time, from Physics to Phenomenology)*, Paris, 2006.
- [9] J. Levine, "Materialism and qualia: the explanatory gap," *Pacific Philosophical Review*, 64, pp. 354-361, 1983.
- [10] D. Wineland, P. Ekstrom, and H. Dehmelt, "Monoelectron Oscillator", *Phys. Rev. Lett.*, 31, 1279, 1973.
- [11] A. Aspect, P. Grangier, and G. Roger, "Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities", *Phys. Rev. Lett.*, 49, 91, 1982.
- [12] S. Gleyzes et al., "Quantum jumps of light recording birth and death of a photon in a cavity," *Nature*, vol.446, pp. 297-300, 2007.
- [13] C. Shannon, "Communication in the presence of noise". *Proceedings of the Institute of Radio Engineers*. 37 (1): 10-21, 1949.
- [14] V. Kotelnikov, <http://ict.open.ac.uk/classics/1.pdf> [retrieved: April, 2022].
- [15] I. Reznikoff, "On a realistic and discrete approach to physical time," HAL open archives, hal-03191083, v1, 2021.
- [16] J. McFadden, "Synchronous firing and its influence on the brain's electromagnetic field," *Journal of Consciousness Studies*, 9, pp. 23-50, 2002.
- [17] B. A. Wallace, *The taboo of Subjectivity: Toward a New Science of Consciousness*, Oxford: Oxford, 2000.
- [18] J. Levine, *Purple Haze: The Puzzle of Consciousness*, Oxford: Oxford, 2001.
- [19] G. Berkeley, *Essay towards a New Theory of Vision*, contents: 43 and 117, 1709.
- [20] B. L. Lancaster, "On the relationship between cognitive models and spiritual maps," *Journal of Consciousness Studies*, 7 (11-12), pp. 231-250, 2000.

Smart Factory Automation for Robotic Production of Satellite Formations and Constellations

Klaus Schilling

Computer Science VII: Robotics and Telematics

University Würzburg
Würzburg, Germany

e-mail: schi@informatik.uni-wuerzburg.de

Abstract—In space industry, the emerging mega-constellations, demanding thousands of satellites per year, dramatically change production approaches for satellites. Here, the transition from traditional human labour-intensive manufacturing of few satellites towards automated, robot-supported, flexible industry 4.0 production systems is progressing. The goal of this contribution is an outline of the production steps supported by automation, robotic manipulators and mobile robots for a flexible flow of materials. Specific requirements concern integrated production and testing capabilities to comply with the challenging space environment and to guarantee the mandatory high quality performance standards.

Keywords-satellite production; intelligent distributed system; industry 4.0; robotics; intelligent manufacturing; flexible manufacturing; integration and testing .

I. INTRODUCTION

The sector of space technology sees currently disruptive changes often described by “New Space” or Space 4.0. The traditional large, multifunctional spacecraft in geostationary orbits are complemented by networks of hundreds of small satellites in Low Earth Orbits (LEO) [1][2]. Application areas address telecommunications (such as Starlink with about 1700 satellites or OneWeb with about 400 satellites launched) and Earth observation (such as Planet with about 400 nano-satellites of only 6 kg in orbit). There are plans for placing several thousand further satellites to provide a continuous service with low latency, in particular for Internet of Things (IoT) applications.

While traditionally less than a hundred satellites per year were manufactured in handwork, now small series production approaches are required to achieve outputs of a few thousand per year. The goal of this contribution is to address factory automation and robotic production processes to realize in a short period of time a sufficient number of satellites in orbit to offer the intended communication services. In parallel, the usual extremely high quality standards in integration and testing are to be guaranteed to support operations of the multi-satellite system with appropriate performance under harsh space environment conditions. Thus, challenges for advanced cyber-manufacturing systems result.

This contribution discusses in Section II the standardized satellite design suitable for automated production. Section III addresses the use of automation in the integration process, while Section IV has emphasis on related quality control and test approaches.

II. SATELLITE SYSTEM DESIGN

The challenges in the transition to automated production start already with design of the satellite suitable for later automated serial production. In contrast to traditional human labor and time intensive satellite manufacturing, now standardization and modularization of the satellite bus and of the satellite subsystems become key elements [3] - [5].

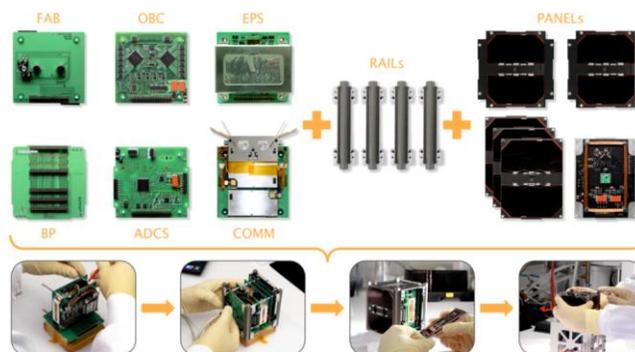


Figure 1. Modular building blocks of the satellite subsystems, supporting efficient integration onto a backplane by standardized electrical interfaces.

A digital twin accompanies the complete satellite life cycle, but in parallel modular hardware building blocks (Figure 1), too. This way, early Hardware-in-the-Loop (HiL) tests can be introduced from the beginning in order to calibrate the digital twin by real hardware performance. Typical satellite system building blocks, like subsystems for on-board data handling, communication, attitude determination and control, energy storage and distribution, are plugged by connectors on the base plate (cf. Figure 1).

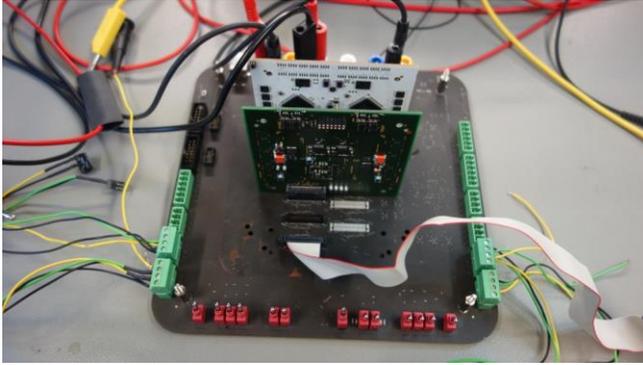


Figure 2. The UNISEC FlatSat development kit providing access for a broad range of test instruments to subsystem hardware, plugged into the base plate.

UNISEC Europe provided an advanced electrical interface definition [6] suitable for very small satellites including CubeSats. All data and power lines are physically placed in the backplane. This replaces the usual harness. Advantages of the proposed bus are robust and rapid development, integration and testing of the satellite as well as simple maintenance, extension and replacement of subsystems in the complete satellite lifecycle, from development (in a flat-sat configuration) to flight model realization (cf. Figure 2). When the satellite is in orbit, the FlatSat serves as Electrical Ground Support Equipment (EGSE), to analyze anomalies observed in satellite operations.

III. AUTOMATION IN SATELLITE INTEGRATION

There is automation already employed at the level of components and subsystems. For example, in Figure 3, the gluing of solar cells on the side panels for power generation is presented. Here, by a pneumatic end-effector, the fragile solar cells have to be placed very accurately on the side panel position to be fixed by adhesive and subsequently electrically contacted. This delicate process is handled by a force-controlled two arm robot. Software continuously



Figure 3. Robotic assembly of a satellite solar array.

monitors the health status of solar cells as well as the status of the adhesive. In particular, if a solar cell breaks during manipulation, quick reaction is necessary to avoid gluing a faulty part and thus producing a defect side panel.

The automation level in satellite integration has to be adapted according to the design complexity and the number of satellites. For traditional production of just small quantities, human labor will do the integration. With increasing quantities, human-robot collaboration increases efficiency and reliability (cf. Figure 4), finally leading to fully automated processing for large quantities (cf. Figure 5). Here, robots with force control are employed to guarantee the safety of the human collaborator.



Figure 4. Human-cooperation at Zentrum for Telematics, Würzburg, to jointly integrate and test satellites. The robot acts as a third arm and contributes high precision motion capabilities.

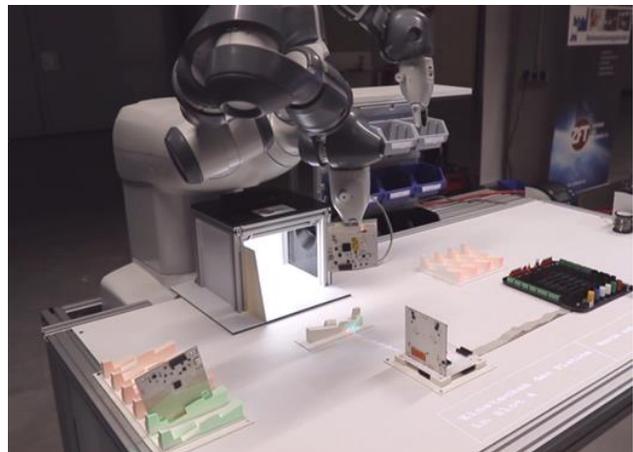


Figure 5. Fully automated integration: here, after optical quality check in the illuminated box, the robot arm moves the board to plug it into the baseplate.

In the fully automated mode (cf. Figure 5), the completed component and subsystem boards are picked by the robot from the storage area, and, after satisfactory quality checks, the integration into the base plate is performed. It uses smart

force control of the robot to enable the very sensitive plug-in of connectors without destroying the baseplate and nevertheless providing a solid connection to tolerate the high vibration level at rocket launch.

IV. INTEGRATED SPACE ENVIRONMENT TESTING

After each integration step tests are applied to assure performance. Often, for more extensive space environment tests, the integrated parts have to be transported from the robotic work cell (cf. Figure 5) to a dedicated test area for assessment of vibration, radiation or motion and pointing capabilities. This transport is autonomously realized by mobile robots to the target facility (cf. Figure 6). Coordination by information exchange via communication links in the production environment is essential between the involved machines. In addition, results are to be transferred to the factory’s process planning system in order to potentially initialize a re-planning process in case of anomalies.

Similar approaches are, therefore, considered in research studies as basis for future in-orbit assembly [8]. The limitations in volume and mass by the rocket shroud often lead to very complicated designs to fold bulky parts, like deployables or arrays, for the delivery into orbit. It is always a critical and failure prone operation to unfold after deployment from the rocket. It would, therefore, be much easier to transfer the components and then finally integrate them by similar robotics to complete satellites in orbit.

V. CONCLUSION

The planned mega-constellations demand thousands of satellites to be produced in a short timeframe. This cannot be achieved by traditional satellite manufacturing. Therefore, modern consumer mass production technologies, like industry 4.0, are analyzed for transfer and adaptation to small series suitable for space environments. An approach based on modularization and standardization of satellite subsystems and components was developed and tested for

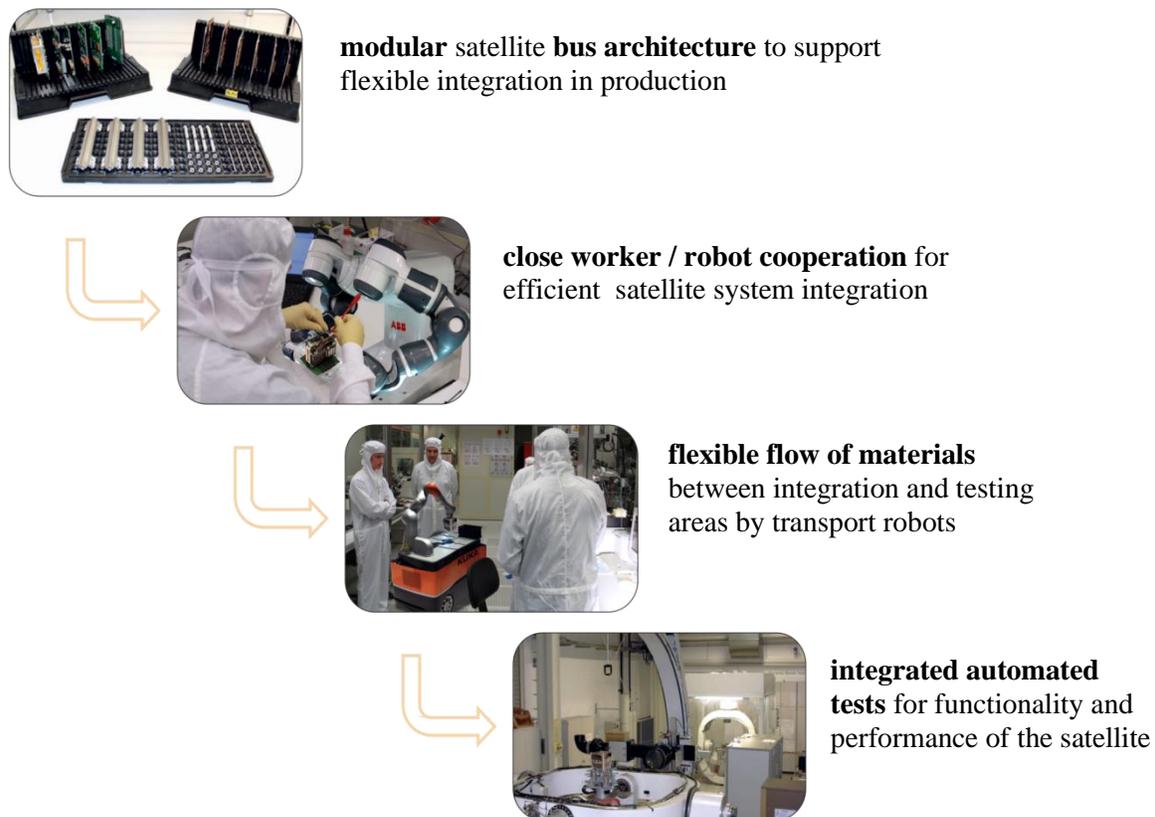


Figure 6. Process flow in satellite integration with integrated space environment testing.

Related advanced robotics and automation approaches were implemented at the Center for Telematics in Würzburg as demonstrator plant, using the example of nano-satellites. Care was taken that the applied methods are scalable to larger satellites, too. Satellites assembled this way were already placed successfully in orbit (NetSat mission launched 2020 [7]).

nano-satellites. It demonstrated the potential of modern automation and robotics technologies in order to increase efficiency in satellite assembly.

This technology demonstrator for integrated assembly and test proved first principles and processes for efficient automated production. Suitability to comply with the challenges of the harsh space environment was

demonstrated in orbit at the NetSat mission launched September 2020. Future developments will now be practiced for production of larger quantities of small satellites for telecommunication and Earth observation applications.

ACKNOWLEDGMENT

The financial support for the project “Adaptive Production” within the “Digital Production” Program of the Bavarian Ministry of Economics is acknowledged, as well as the contributions from all motivated collaborators in this project.

REFERENCES

- [1] Acatech, Recommendations for Implementing the Strategic Initiative Industry 4.0 - Final report of the Industrie 4.0 working group, April 2013.
<https://en.acatech.de/publication/recommendations-for-implementing-the-strategic-initiative-industrie-4-0-final-report-of-the-industrie-4-0-working-group/> [accessed May 2022]
- [2] Euroconsult, „Prospects for the small satellite market, 7th edition“, 2021.
<https://digital-platform.euroconsult-ec.com/product/prospects-for-the-small-satellite-market/> [accessed May 2022]
- [3] E. Hand, “Startup Liftoff”, *Science* Vol. 348, Issue 6231, pp. 172 – 177, 2015.
- [4] N. Jones, “Mini satellites prove their scientific power”, *Nature* 508, pp. 300 – 301, 2014.
- [5] K. Schilling and M. Dowling, “Disruptive Technology potential of Small Satellites”, *The 4S Conference Proceedings*, paper 320, 2016.
- [6] UNISEC Europe; CubeSat Subsystem Interface Definition, 2015.
<http://unisec-europe.eu/standards/bus/> [accessed May 2022]
- [7] J. Scharnagl, F. Kempf, S. Dombrowski, and K. Schilling, “NetSat – Challenges of a Formation Composed of 4 Nano-Satellites”, *72th International Astronautical Congress, Dubai*, IAC-21-B4.7.8, 2021
- [8] F. Kempf et al., “AI-In-Orbit-Factory – AI approaches for adaptive robotic in-orbit manufacturing of modular satellites”, *72nd International Astronautical Congress, Dubai*, IAC-21- D1.1.2, 2021.

A Deep Learning based Unoccupied Parking Space Detection Method for City Lots

Hamid Reza Tohidypour¹, Yixiao Wang¹, Panos Nasiopoulos¹, and Mahsa T. Pourazad^{1,2}

¹Dept. of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada

²TELUS Communications Inc, Vancouver, Canada

e-mail: {htohidyp, yixiaow, panosn, pourazad}@ece.ubc.ca, mahsa.pourazad@gmail.com

Abstract— Nowadays, finding a vacant parking space in populated metropolitan cities is a challenging task, leading to serious traffic congestion with environmental and productivity ramifications. Although many different systems have been proposed and tested for unsupervised parking lot space detection over the years, they have been proven to be either impractical or costly to maintain. In this paper, we propose a deep learning based approach that uses the video captured by a vehicle mounted camera to accurately detect and count the number of available parking spaces in real-time. Our system achieves an impressive average detection accuracy of 90.59% for unoccupied spaces and 95.66% for the occupied spaces.

Keywords- deep learning; city lots; object recognition; YOLOv4.

I. INTRODUCTION

Properly managing parking lots and displaying available parking information to the drivers before entering the parking lot has been a challenge. Many different systems have been employed over the years for unsupervised parking space detection for open lots, including counting the number of cars entering and exiting the parking lot or using a network of ground- or ceiling-mounted occupancy sensors [1]. Although the latter is a welcome solution, it turns out to be a costly one as sensor maintenance proved to be an expensive proposition. In this paper, we propose a deep learning based approach that uses the video captured by a vehicle mounted camera to accurately detect and count the number of available parking spots in real-time.

The rest of this paper is organized as follows. In Section II we give a detailed description of our data collection, the labeling approach, and the deep learning network architecture that we used. Section III presents the performance evaluation of our method and discusses the results. Finally, Section IV concludes our paper.

II. OUR PROPOSED METHOD

Our objective is to design a deep learning based real-time occupancy detection system, using a video feed from security and parking inspection vehicle-mounted cameras. To this end, our first task was to capture a comprehensive and representative dataset to use for training and validating our deep learning network. The second task involves the design of an effective labelling scheme that will allow our network to accurately detect the available parking spots in a parking lot.



Figure 1. a) image captured by GoPro Hero4 vs b) image captured by AKASO EK7000.

A. Data Collection

In generating an appropriate dataset, we first had to determine what type of commercial camera may be used that meets our requirements such as frame rate and resolution. We compared two cameras namely the AKASO EK7000 and the GoPro Hero4 action cameras. It is important to note that the cameras are mounted on top of a moving vehicle. Figure 1 provides a visual comparison between frames captured by the GoPro Hero4 and AKASO EK7000 cameras. As it can be seen, motion blurring is observed in the GoPro Hero4 footage, with AKASO EK7000 yielding consistently better clarity and less motion blurring for our application.

In our implementation, an AKASO EK700 camera was mounted on top of a moving vehicle. Since the angle and distance of the camera from the parking spaces may affect the shape of the objects of interest in the frame, we tried to cover all practical scenarios of distances and angles. The former was achieved by capturing data from multiple parking lots around the city while the latter by using two vehicles of different sizes, a sedan and an SUV. In addition, we tried to cover a huge variety of parking scenarios, with different types of cars, cars that were not “properly” parked and parking lots with different ways of indicating parking spaces. This huge variety of data aims at helping our network to, on one hand, avoid overfitting and on the other to accurately detect the availability of parking spots in unique and unknown situations. A large number of videos was captured and converted to single frames for labelling and training the network. To avoid overfitting from using similar frames, we chose only the frames with some significant difference. A total of 4000 unique frames was used for training our network.

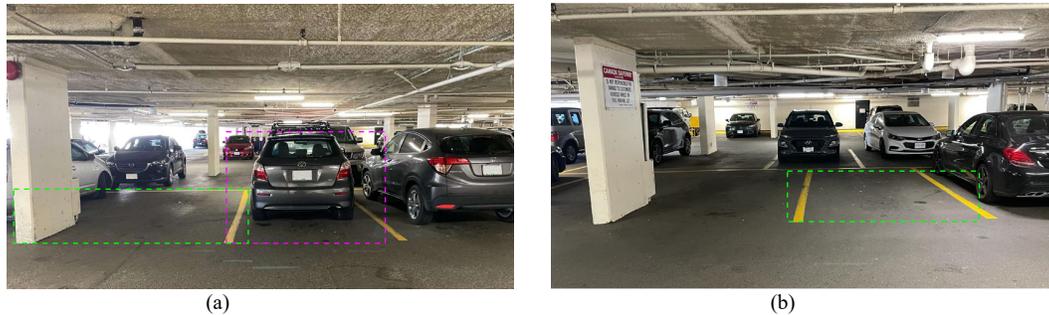


Figure 2. In both frames the car travels from right to left. a) Only the fully visible occupied space is labeled with magenta dashed rectangles and one unoccupied parking space, which is labeled with the green dashed rectangle. b) A frame with two unoccupied spots and one occupied space, which is not fully visible to the camera.

B. Labelling

Our main objective is to design a deep learning based approach that detects unoccupied parking spots and counts the uniquely available parking spaces. Assigning the proper label classes is one of the challenges and a key factor in determining the success of our proposed approach. It is worth noting here that the only difference between unoccupied and occupied parking spaces is the presence of vehicles. For this reason, in order to facilitate the training process, we decided to add a class for occupied parking spots as well, thus assigning each frame labels and bounding boxes for occupied and unoccupied parking spaces.

In the case of object recognition, it is common practice to label all the objects of interest that are present in the scene. In our implementation, labeling of parking spaces has to follow a predetermined rule that will allow the deep learning network to learn to track the spots from frame to frame and thus accurately count the number of unoccupied spots. It is worth noting that a key feature of a parking space is the presence of two lines that indicate the borders of the spot. On the other hand, the difference between an unoccupied and an occupied parking space is the presence of a car in the latter case. Given the above observations, we decided to have two labels, occupied and unoccupied parking spaces. Since in this project, we want to count the number of unique unoccupied parking spots, in every frame we considered to label only the first unoccupied spot we encounter, if it exists. In the case of occupied spots, we label in each frame all the spaces that we clearly see, i.e., almost perpendicular to the camera and we see the car. Figure 2 shows an example of two labeled frames. In both frames the car is traveling from right to left. In Figure 2 (a) we observe two occupied spaces. In this case, only the occupied space that is fully visible to the camera is labeled with the bounding box shown by a magenta dotted line and the unoccupied space labeled by a bounding box shown with a green dotted line. Figure 2 (b) depicts one occupied space, which is not fully visible to camera, and two unoccupied spaces. In this case, only the first unoccupied space is labeled with the bounding box (green dotted line). As the vehicle-mounted cameras move forward at an expected speed of vehicles driving in a parking lot (15kmph-20kmph), our counting algorithm described in

detail is Section III considers only the next unique frame and thus unoccupied space.

C. Deep Learning Network

Our goal here is to use a Deep Learning network architecture that is designed for object detection and offers real-time performance. Although there are many widely used neural networks for object recognition, ranging from Faster R-CNN to Mask-RCNN and YOLO (You Only Look Once), the performance of the 4th generation of the latter seems to meet all our requirements [2] [3] [4].

YOLO is a family of one-stage object detectors that are accurate and fast and their architecture allows them to be trained much easier than the other state-of-the-art networks.

YOLOv4 introduces a set of advanced features, such as spatial attention modules, a new activation function, new data augmentation schemes and Self-Adversarial Training, which significantly improve the performance of the network compared to previous versions.

Figure 3 depicts a comparison of YOLOv4 with other state-of-the-art object detection networks, clearly showing that it is much more accurate than YOLOv3 and much faster than EfficientDet with comparable performance [4].

To make sure that we chose the correct network architecture, we trained both the EfficientDet and YOLOv4 using our dataset. YOLOv4 started converging much faster, using only one CPU and a smaller mini-batch size, while EfficientDet could not reach similar performance, mainly due to memory requirements and possible overfitting due to the limited size of our dataset. For these reasons, we decided to use YOLOv4 for our task.

III. PERFORMANCE AND EVALUATIONS

We trained YOLOv4 using a state-of-the-art computing cluster [5]. Our dataset consisted of a large number of video sequences, which were converted to 4000 unique frames that were used for training and validation (80% and 20%, respectively). Note that the suggested default settings of the YOLOv4 for batch size is 64 and the learning rate is 0.001. In addition, the recommended threshold for intersection over union for the mean average precision (mAP) is 0.5, while the resolution of input images is 608x608 pixels, i.e., inputs are downscaled to that size by YOLOv4. When we used these

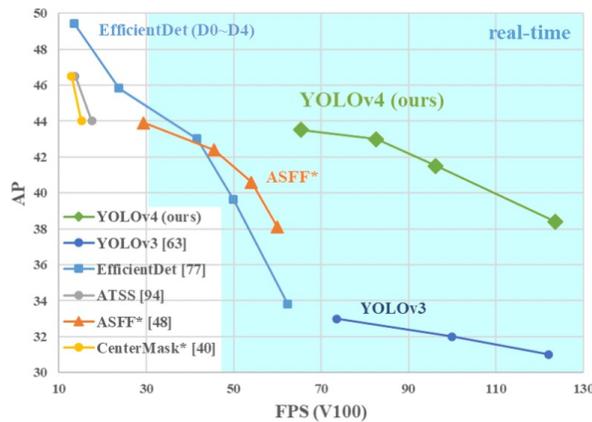


Figure 3. YOLOv4 is much more accurate than YOLOv3 and much faster than EfficientDet with comparable performance [2].

recommended settings for training this network, the mAP reached 94% at the early stages of training, gradually dropping to mAP of 90.7%, showing that the hyperparameters are not well optimized for our task. To address this issue, we investigated using different hyperparameters for YOLOv4, in which we reduced the learning rate to 0.0001 while increasing the batch size to 128. Total training time was approximately 48 hours. Figure 4 shows the loss vs iteration number for the training stage and the mAP vs iteration number for these new settings. As can be seen in the figure, the loss plateaus during training, eventually leading to 96% accuracy for the best weight sets, which we use for our testing model.

In order to test our model, we used a MacBook Pro laptop with M1 chip with 8-core CPU, 8-core GPU and 16-core Neural Engine. The laptop proved to be very powerful and run our model in real-time using a webcam feed, performing real-time predictions at 30fps.

Our model’s prediction accuracy is measured based on its ability to classify the unoccupied and the occupied parking

spots classes. In addition, a counting algorithm was developed that used the occupancy detection model to count the number of distinct parking stalls.

Several testing conditions related with the parking lot may affect the occupancy confidence levels of our detection model. To eliminate many unknown variables, we tried to design our model to be as robust as possible to such conditions. As discussed before, we consider vehicle speeds ranging from 15km/h to 20km/h, a range acceptable for security and inspection cars traveling in a parking lot. Our labeling process and the architecture of our network allow us to account for this range of speeds and the fact that the speed of the vehicle most likely will vary during different rounds in the parking lot. As the camera is mounted on the vehicle, the speed at which it moves through the parking lot determines how many frames are captured per area/parking space. The above is an accurate statement, as the relative distance of the camera from the parking space stays almost the same and our tests showed that small variations do not significantly affect the number of frames per area.

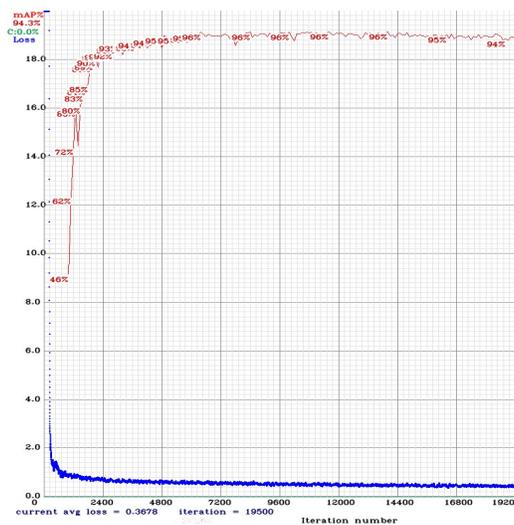


Figure 4. The loss and mAP@0.5 plot for YOLOv4. Total training time was approximately 48 hours. The mAP@0.5 reaches 94.3%, with the best mAP@0.5 for 96%. The average loss value floats around 0.3678.

Prediction accuracy is also determined by the prediction threshold. With each prediction performed by the model, the model attaches a confidence level which indicates how likely a class object exists within the bounding box. Using a threshold (Thr) level allow us to only consider predictions with a confidence level higher or equal to the threshold level; those are the ones that will be displayed with bounding boxes.

Performance evaluation of our deep learning trained model is based primarily on its accuracy in detecting unoccupied and occupied parking spaces, using several testing scenarios. The scenarios of main interest consist of unoccupied and occupied regular and angled spaces, reserved parking spaces and handicap spaces. Exceptional cases include the presence of a wall on one side of a parking spot instead of a parking line, cars parked over parking lines, obscure parking lines due to erosion of paint, corner spots, poorly visible parking spots due to rain, and parking spots using special color lines instead of the regular white parking lines.

Our test results are summarized in Table I, which shows the detection accuracy (AP) of the two classes, occupied and unoccupied. Real-time predictions were done on the laptop positioned inside the car that was receiving a 30fps video feed from a camera mounted on the top of the vehicle. We observe that for confidence threshold 30% the probability of predicting an unoccupied parking space is 90.59% while for the probability of predicting occupied parking spaces is 95.66%, while for confidence threshold of 50% the corresponding predictions precisions are changed to 88.86% for the unoccupied spaces and 95.52% for the occupied spaces. This means that our network mostly assigns probabilities greater than 50% to the detected classes. However, in order to avoid missing any unoccupied spaces, we decided to use the confidence threshold of 30%.

The fact that occupied parking spaces may be identified by both the presence of lines and cars, which in terms of a neural network means more features, leads to a slightly higher prediction accuracy for the occupied parking spaces than the unoccupied ones.

Our unoccupied parking space counting algorithm is built on the premise that our model always detects the first parking space and only one unoccupied space in each frame. The car speed is used for determining which frame will be considered next, making sure that we will “see” a new parking space introduced in the next frame. For instance, at 15 km/h the car will cover 4 meters in 1 second. At the same time, our camera captures 30fps. Since our measurements have shown that a parking space in a city parking lot is

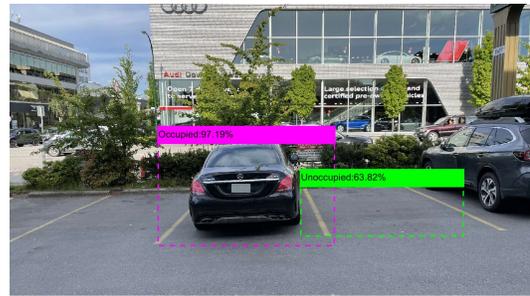


Figure 5. An example of test images with the predicted bounding boxes, labels, and the probabilities assigned to them.

approximately 3.5 meters wide, we concluded that we should process every 22nd frame so we ensure that we have a new parking space introduced to the frame. This calculation is based on the assumption that the parking lot is normal size and the car camera is facing the side of the passenger, which is closer to the traveling car. Such a scenario is shown in Figure 5. In case that the parking lot is larger in size or we want to position the camera to view the cars on the driver’s side, thus increasing the distance between the camera and the parking spaces, the algorithm may be adjusted to consider a different number of frames. For example, if the distance is double as shown in Figure 6, a new parking space will be introduced every 16 frames. This example shows the challenging case of a rainy scene with wet asphalt. As it can be seen in Figure 6, our model was able to correctly detect the parking space and the occupied space next to it. The probability that was assigned to the unoccupied space was 37.13%, which is less than the probability that was assigned to unoccupied space in Figure 4, due to the weather conditions. However, this shows that assigning confidence threshold of 30% makes our model robust to the ambient conditions.

IV. CONCLUSIONS

In this paper, we introduced a deep learning based detection method for unoccupied parking spaces in a city parking lot. We developed a unique labeling scheme, which allowed us to detect one new unoccupied parking spot in a



Figure 6. An example test image for the case that distance is two times larger than the case of Figure 5.

TABLE I. TEST RESULTS FOR REAL TIME SCENARIO.

| Parking Status | Average Precision | |
|----------------|-------------------|-----------|
| | Thr = 30% | Thr = 50% |
| Unoccupied | 90.59% | 88.86% |
| Occupied | 95.66% | 95.52% |

frame and thus accurately count the total number of vacant spaces in the parking lot. We used the YOLOv4 network and we achieved a detection accuracy of 90.59% for unoccupied spaces and 95.66% for the occupied spaces. As our training and validation dataset was only 4000 frames due to privacy issues, we are confident that this network can achieve even higher accuracy with a larger dataset and the same labeling scheme.

ACKNOWLEDGMENT

We would like to acknowledge the contribution of C. Yao, M. I. Islam, R. Chu, S. Ly, and B. Wang for data collection and initial work on the subject. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC – PG 11R12450), and TELUS (PG 11R10321). This research was enabled in part by support provided by WestGrid (www.westgrid.ca) and Compute Canada (www.computecanada.ca).

REFERENCES

- [1] V. Paidi, H. Fleyeh, J. Håkansson, and R. G. Nyberg, “Smart parking sensors, technologies and applications for open parking lots: a review,” *IET Intelligent Transport Systems*, vol. 12, no. 8, pp. 735 – 741, May 2018.
- [2] A. Bochkovskiy, C.-Y. Wang, and H.-Y. Mark Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” arXiv preprint arXiv:2004.10934, 2020.
- [3] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *The 28th International Conference on Neural Information Processing Systems*, vol. 1, pp. 91-99, December 2015.
- [4] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” *Proc. IEEE International Conference on Computer Vision*, IEEE Press, pp. 2980-2988, Oct. 2017, doi: 10.1109/ICCV.2017.322.
- [5] Compute Canada state-of-the-art advanced research computing network. Available from: <https://www.computecanada.ca> [retrieved: April 5, 2022].

Semantic Patterns to Structure TimeFrames in Text

Nour Matta

IT and Digital Society Laboratory
University of Technology of Troyes
Troyes, France
email: Nour.matta@utt.fr

Nada Matta

IT and Digital Society Laboratory
University of Technology of Troyes
Troyes, France
email: Nada.matta@utt.fr

Nicolas Declercq

Namkin
Troyes, France
email: n.declercq@namkin.fr

Agata Marcante

Research & Development
Namkin
Troyes, France
email: a.marcante@namkin.fr

Abstract— Event ordering is a very important task in the event extraction field since any analysis of the causality and impacts of a specific action or a change requires consideration of temporality and ordering. Many pattern-based approaches or machine learning approaches work on identifying the events in the text and creating relationships between them. In this paper, we present a novel approach based on timeframes, that will enable distinction between multiple timeframes in a text, when available, and grouping events within these timeframes.

Keywords-Timeframe; Event Extraction; Event Ordering; Natural Language Processing.

I. INTRODUCTION

Event extraction is one of the most important tasks of Information Extraction through Natural Language Processing [25]. It enables the extraction of events in text and aims to identify the different participants and attributes of the extracted events. Some examples of the extracted information can be the cause, place, time, means, or goal, that can be identified through dependency analysis [28]. Moreover, evaluating the influence of a particular event or a specific action requires an account of temporality [27]. In the traditional event extraction, available approaches are very performant when it comes to the analysis of single sentences. Some approaches can support complex sentences. But even though models aim to extract events from a “text” and create temporal relations between them, the performance lacks and soon the extracted information easily becomes unreadable or, from a temporal relation point of view, inaccurate [8]. Furthermore, when focusing on the temporality event extraction, many approaches focused on ordering the multiple events mentioned one by one and creating relationships [12], [26] without considering the fact that multiple ‘processes’ can be part of a preparatory stage of a single event. In this paper, we introduce the use of timeframes, an approach used for the time analysis in different domains, to improve the temporal relation made between events in a text.

It is important to note that within the same text, multiple timeframes can be identified, and multiple time references can be used. A small example would be a news report about a company announcing the launching of a new product. We have the time when the announcement was made, the timeframe within the announcement (such as the date of the launching), and the time of the publication of the news. Another example would be in a narrative text in which the author talks about multiple events while going back and forth in time. Our main goal is to identify the events in a text, create temporal relations between them and identify the different timeframes if there are multiple ones. We aim to assign each event to its timeframe enabling improved readability of the extracted event, their temporal relation, and finally their interpretation.

In Section 2, we will start by defining what timeframes are and how they are used for time analysis. We will also present the different conceptualizations of events in linguistics. In Section 3, the related work, we will go through the different event extraction approaches before presenting the timeframe approach along with part of the semantic pattern identified.

II. TIMEFRAMES AND EVENTS

This section is divided into two main parts: the timeframes and the events. For the timeframe, we will go through the analysis of temporality in fields other than text mining and show how those conceptualizations can be helpful in the analysis of temporality in text. As for the events, we will go through their definition in the event extraction field and how it is viewed from a linguistic point of view.

A. Timeframes

A timeframe is a certain period of time in which an event should happen or has already happened [3]. This leads us to question the meaning of time. In philosophy [19], the platonist understanding of time is segregated from the relationist definition. Platonists picture time as an “empty container” of events that exists regardless of whether

anything is placed in it. In this perspective, platonists consider that it is possible that changes in the universe can cease to exist for a certain period. On the other hand, relationists view time as a set of events and the temporal relationship between these events. While dealing with event extraction and their temporal relationship, the relationist understanding is used.

The study of the temporal relation between actions, events, states, and their influences is applicable in different domains other than event extraction. In their study on temporality in video games, Zagal et al. [27] distinguished multiple types of games: the ones with game time being equivalent to the real-world time, the ones in which action can speed up or skip time, the ones where specific action triggers events of a specific duration, and finally the ones where certain events occur without affecting the game time as if time had stopped. To analyze the temporality for each game type, Zagal et al. [27] defined timeframes, creating relations between those timeframes and between events within the timeframe and coordinating them. Reflecting on that approach, from a textual perspective, the authors also set the duration to specific events as shown in “1)”, which can make flashbacks “2)” and flash-forwards “3)”. They can also skip time “4)” and even focus on a specific event or describe elements making the time indirectly stop “5)”.

- 1) *John ran for an hour.*
- 2) *Henry was looking at the photo. He took it a few years back when he was in New York.*
- 3) *John is preparing his luggage; he will be leaving in the morning.*
- 4) *Five years later, Henry went back to New York.*
- 5) *John looked through the window for a few seconds. It was a rainy day; people were walking while holding their umbrellas. He went to his desk.*

Distinguishing the different timeframes and specifying the events that happened in each frame enables the focus on specific events based on their occurrence time and aims to improve coordination between multiple timeframes in the text. However, in the event extraction field, events are ordered one by one without having a more global representation. Some of the concepts that must be considered while dealing with temporality are the duration, the time point, calendar, narrative time, timeline, countdown, and temporal relation [3], [27]. Each of these concepts plays a specific role in the pattern and the extracted knowledge. The use of timeframe also enables the consideration of the release date of the text as a timeframe on its own in order to improve topic tracking and event follow-up.

B. Events

In the event extraction field and the event-based decision systems, events are usually defined as happenings or changes that occurred in a specific interval of time. They can be associated with the change of states (canceled, ongoing, recently done, past or future plans) and can have multiple occurrences [17], [25].

Other than Natural Language Processing, linguists also worked on defining what an event is, distinguishing it from a

state, and partitioning it onto atomic and extended events. Using the tense of the verb, the duration, and time reference along with temporal connectors, some set of rules and patterns are proposed. Vendler [23] was one of the first to work on defining the concept of event in linguistics, while working on verbs and tenses, he first identified the tense as the location of a happening in the time (past, present, or future) and its aspect which refers to the state of an event (completed, ongoing or interrupted). Later on, he defined “Eventualities” [23] as a concept that groups the state and non-state. Some particularities of each group were identified:

- 6) *Jack was ill on Sunday.*
- 7) *Jack wrote a letter on Sunday.*

“6)” is an example of a state, in which we cannot determine if the state “ill” started before or during the “Sunday” and if is stopped during “Sunday” or after. While the non-state “wrote” started on ended Sunday. And comparing the duration of “was ill” and wrote, we can presume that “wrote” has a shorter duration than “was ill”.

Their conceptualization of eventualities goes as follows: non-state was divided into Activities and Events; activity refers to actions that had a duration but with no endpoint or consequent state while events have a quantification or an ending result:

- 8) *Alex ran.*
- 9) *Alex ran to the store.*
- 10) *Alex ran a mile.*

“8)” is considered an activity while “9)” and “10)” are events. Events are then distinguished [24]. Where an accomplishment is considered to have a duration and accept the progressive (continuous tense) while achievement is strange in progressiveness. It is important to note that Kamp highlighted the ambiguity between those concepts, starting with the very first division between distinguishing a state and a nonstate.

Using the conceptualization made by Vendler, Moens et al. [17] defined another conceptualization. Eventualities are divided into States and Events. And they considered two dimensions for distinguishing events: the duration and the consequence. For the duration, they considered events as Atomic or Extended. Extended Events have a notion of duration. For the consequence, they started by defining the term “culmination” as an event that has a consequence, a change of state. A “nucleus”, as shown in Figure 1, is the combination of a preparatory process of a culmination, the culmination, and the consequent state. If we consider the example “9)”, “Alex ran to the store”, we can regard it as a culmination, in which running is a preparatory process to the culmination of arriving at the store. The consequent state is “being in the store”.



Figure 1. Moens et al. nucleus definition [17].

Figure 2 presents the 4 subcategories of events. An Atomic Event with no consequence is considered as a point, for instance, “He hiccupped”. A culmination is an atomic event with a consequence [17]. An extended event is a process and is considered a culminated process if it has a consequence. It is important to highlight that many elements were used to distinguish the different categories of events. This work and the pattern identified in it are essential for our approach especially the use of a nucleus. When using the timeframe approach, identifying the culminations in a text and all the processes, the preparatory stage, and the consequent state are one of our goals. The slice difference in our approach is that we intend to associate different events on the nucleus timeline.

| | Atomic Event | Extended Event |
|---------------|--------------|--------------------|
| - Consequence | Point | Process |
| + Consequence | Culmination | Culminated Process |

Figure 2. Moens et al. event conceptualization [17].

The pattern identified by Moens et al. will be used and associated with other patterns to enable the representation of events extracted for each timeframe. When trying to identify events and states, the use of adverbs, the tense of the verb, and the use of semantic dependencies, such as the verbs’ objects, were used for identifying the categories. The same verb can be considered a point, a process, a culmination of a culminated process depending on its use.

It is important to note that in linguistics, verbs tend to be classified as states and events. Adjectives are considered states of the elements they describe. In the event extraction field, nouns are also identified as events depending on their context, for example:

11) *Two years after his graduation, John moved to New York.*

In this sentence, the noun “graduation” is considered as an event. In order to manage these types of events, a pattern concerning nouns was added. If a noun is a temporal reference, in this case, “after graduation”, then this noun is an event. Several studies define principles to identify and extract events along with their arguments; we summarized them in the related work.

III. RELATED WORK

This section will be partitioned as follows: we will start by going through event extraction techniques and more importantly Event Ordering. Then we will go through different research works that addressed temporality aspects in the temporality recognition techniques.

A. Event Extraction and Even Ordering

There are two main event extraction types: the closed-domain and the open-domain [26]. The closed-domain event extraction refers to the detection of specific events of

interest, for example, the merger of two companies. In this case, the information related to the event is predefined. This approach is usually used for event mention detection, event trigger, and event argument [12]. Some classify the role of each argument that is specific to the event of interest. As for the open domain, the search of events is not bounded to specific events and aims to detect all sorts of events within a text and later on cluster texts based on similar events detected [8]. This approach is usually used for story segmentation, first story detection, topic detection, etc. In this paper, we focus on the open domain event extraction since applying the approach to a specific domain is out of our scope. In order to extract or detect events, multiple approaches are available from pattern matching to machine learning [13], deep learning [10], semi-supervised learning [7], and unsupervised learning [26]. This depends on the type of events of interest, if there are pre-trained models, and the purposes of the extraction.

Event Ordering is a branch in event extraction that focuses on extracting events and creating temporal relations between them [15]. Multiple annotated datasets are available to train models, the most popular being the TimeBank-Dense corpus [4]. This corpus has three types of relations, the intra-sentence, the cross sentence, and the document creation time. Note that one approach was considered as a “context-aware” model for using all three types of relations [16]. Temporal label dependencies and constraints are used to improve relations between events [1]. Some worked on the linguistic and syntactic rules, such as Leeuwenberg et al. [11] or Laokulrat et al. [10]. Please note that, in this paper, we will be using the defined constraint, rules, and linguistic features to enrich our approach. The relation between a timeframe and events will be added with consideration of culminations and constraints on culminations.

B. Temporality Recognition Techniques

In the Question Answering field, temporal analysis is a must for determining if an answer to a question will change throughout time or not. In their work, Pal et al. [20] identified multiple classes of information temporality: short duration, medium duration, long duration, and permanent. They tried classifying the question/answer under those categories but ended up grouping the short term and medium term together and long term and permanent together. It enabled distinguishing between “Who won the competition X in 2022?” and “Who won the last competition?”. One of the questions will have permanent information and the other will change throughout the years. It is important to consider these types of classifications to identify information that is true regardless of the timeframe of the text and relations that are relative to the timeframe of the text. Recent work focuses on identifying the attention in complex questions and the use of multiple sentences that contain the answer [5]. Note that in their work Kwiatkowski et al. [9] mentioned descriptive sentences or informative sentences, in which information is given without a particular event being mentioned.

Temporality plays a very important part in social science and social discourse analysis [6]. Coordination between different events from multiple resources is also used when

clustering news and following up on events. Sources vary between news and social media posts, such as tweets [18]. It is also essential to consider time relations when analyzing the influence of social media and the media in general on social events, such as protests and violence and study the sentiments behind it [22].

The question answering field provided a very important aspect to consider when extracting events and information. Completed Events and states with a specific date tend to be permanent information while unfinished events and events with reference to the text temporality tend to be true in a specific timeframe. Coordination of events between multiple texts will be considered in our approach and will be based on the timeframe concept. Our approach introduces the use of multiple types of timeframes and how to extract them. We will be using several models and patterns already provided in order to optimize the model's performance.

IV. TIMEFRAME APPROACH

For the extraction of timeframes that will be used to improve the temporal relation analysis between events, we identified three types of timeframes, (1) the Publication Timeframe, (2) the Narrative Timeframe, and (3) the Spoken Timeframe. Those timeframes were inspired by the identified timeframes for temporal analysis in video games with adaptation to the text constraints [27]. The Publication Timeframe reflects the publication date or year of the analyzed text. The Narrative Timeframe is the timeframe of the events happening in the text; we may find multiple Narrative Timeframes in a single document. Finally, the Spoken Timeframe is a particular type of timeframe that may not always appear in a text. It is used when an announcement, a speech, or a dialogue is present. The events and information that are mentioned in that context will be analyzed in their own timeframe in order to reduce event relationship complexity. The timeframe will consist of two main parts: (1) the text belonging to the timeframe, and (2) the extracted information related to it.

A. Publication Timeframe Extraction

All text document have by default a Publication Timeframe and a Narrative Timeframe. To identify the publication date, the type of text affects the extraction. If a post on social media is being analyzed then, the date is usually available as metadata to the text. When dealing with online news, most publishers put the date at the beginning of the text. Considering the presence of the title, we will check the first three sentences, for the presence of dates using Named Entity Recognition. If no dates were found, the last sentence will be checked. In case a sentence was identified as the publication date, it will be extracted from the document in order to avoid confusion with the rest of the text. The date will be set in the information field of the timeframe. Figure 3 provides the Publication Timeframe extraction function. It takes two elements as input: a text, and the patterns that identify the publication date. The returned list contains two elements: the Publication Timeframe and the text. The text is returned since it is modified in case the pattern was found in a sentence.

```

1 Extraction_Publication(Single_Text, pub_pattern):
2   sent = split_sentences(single_Text)
3   tf_pub = [], []
4   to_check = [sent[0], sent[1], sent[2], sent[-1]]
5   for element in to_check:
6     identified = check_pattern(pub_pattern, element)
7     if identified:
8       timeframe_pub = [[date], [element]]
9       remove element from Single_Text
10      return [timeframe_pub, Single_Text]
11  return [], Single_Text

```

Figure 3. Publication Timeframe Extraction.

Table 1 provides some of the patterns used to identify the Publication Timeframe. Please note that for the first two patterns, their presence in the sentence is enough while for the last two, they must be alone in the sentence to be considered a sign of Publication Timeframe.

TABLE I. SOME OF THE PATTERN USED TO DETECT PUBLICATION TIMEFRAMES

| pub_patterns |
|---|
| 'Updated' + <date> |
| 'Published' + <date> |
| <number> + [hours, days, months, years] + 'ago' |
| <date> |

B. Spoken Timeframe Extraction

```

1 Extraction_Spoken(Single_Text, say_pattern):
2   list_sentence = split_in_sentences(single_Text)
3   tf_speech_element = [], []
4   tfs_speech = []
5   id = 0
6   before = false
7   for sentence in list_sentence:
8     identified = check_say_pattern in sentence
9     if identified and not before:
10      before = true
11      tf_speech_element = [id, [sentence]]
12      replace(sentence, Single_Text, "tf_speech_" + id)
13    else:
14      if identified and before:
15        add sentence to tf_speech_element[2]
16        remove sentence from Single_Text
17      else:
18        if before:
19          before = false
20          add tf_speech_element to tfs_speech
21          id = id + 1
22          tfs_speech = [], []
23  return [tfs_speech, Single_Text]

```

Figure 4. Spoken Timeframe Extraction.

This timeframe will be treated before the Narrative Time. The search for verbs that reflect speaking and punctuation that are proper to dialogue will be the main task. If nothing is identified, we skip to the next stage, else a spoken frame will be created. If the "spoken" elements are all available in successive sentences, they will all be extracted and set in a single Spoken Timeframe. If multiple sentences have 'spoken' elements but are not successive, a Spoken

Timeframe should be created for each nonconsecutive part. But in order to enable relations between the Narrative Timeframe and the Spoken Timeframe, identification will be assigned to each extracted Spoken Timeframe, and the extracted sentences will be replaced by the Spoken Timeframe Identification. If any dates are mentioned, they can be added to the information field of the timeframe. Note the tense in the Spoken Timeframe reflects a relationship between the Spoken and Narrative Timeframe it belongs to, so if a unique tense is identified, a relation between the Spoken and the Narrative Timeframe will be identified. For example, if future tense is identified in the “spoken” element, then the relationship will most probably be “after”. To keep track of this relationship, the relation if available will be added with the timeframe identification.

Figure 4 provides the Spoken Timeframe extraction function. It takes two elements as input: a text, and the patterns that identify the speaking patterns. The returned list contains two elements: the Spoken Timeframe list and the text. The Spoken Timeframe list contains all the Spoken Timeframes identified in the text. Each one contains an identification that distinguishes different segments in which the patterns were identified along with the sentences. Note that if consecutive sentences contain the patterns, then they will be grouped in the same timeframe. The return text is the remaining text with the identifications of the Spoken Timeframes.

We used a single pattern to identify the presence of a direct speech in a sentence. First, some direct speech may contain multiple sentences between quotation mark which reduced the accuracy of the dependency parsing. This is why, during the pattern matching phase, any text between quotation was replaced by “” and we analyzed the dependencies of the quotations. If the quotation mark is the object of the verb in the sentence, then we consider that the current sentence belongs to a Spoken Timeframe.

C. Narrative Timeframe Extraction

The starting point of the Narrative Timeframe is having an empty information field and the whole text inside of it. The purpose of using multiple timeframes is to distinguish between current time in a text and in case a flashback is mentioned, or flash-forward is mentioned, the information should be treated accordingly. Using the VerbNet parser [2], we detect any temporal relation. We associate a change in the timeframe when the relationship is not related to a specific event. For example, “before going to bed” is related to the event “go to bed” while “a few years ago” is a temporal relation with the current timeframe. We also consider “later that day” or “later that year” elements within the same timeframe.

In this section, we will present the elements that trigger the creation of a new Narrative Timeframe. The temporal relationship elements that will create a new Narrative Timeframe are: “a few years later”, “(number) years later”. The same goes for “months” and “days” instead of “years” and “ago” instead of “later”. Dates are relatively important; if a date is mentioned, it will be assigned as information about the timeframe. If no dates are mentioned, temporal

relations that start with ‘this’ for example, ‘this year’, ‘this month’, ‘today’, will be considered as time information of the timeframe. If multiple dates are separately mentioned, each will be assigned a timeframe.

```

1 Extraction_Narrative(Single_Text, narrative_pattern, tense_patterns):
2   list_sentence = split_in_sentences(single_Text)
3   tf_nar_element = ["tf_nar_0", [], []]
4   tfs_narrative = []
5   id = 0
6   tocheck = 0
7   for sentence in list_sentence:
8     identified = check_narrative_pattern in sentence
9     if tocheck == 0:
10      if not identified :
11       add sentence to tf_nar_element[1]
12      else:
13       dominant_tense = check_tense(tf_nar_element)
14       id=id+1
15       add "tf_nar_"+id to tf_nar_element[1]
16       add tf_nar_element to tfs_narrative
17       sentence_tense = check_tense(sentence,tense_patterns)
18       tf_nar_element = ["tf_nar_"+id,[sentence],[sentence_tense]]
19       if (dominant_tense != sentence_tense):
20        tocheck = 1
21      else:
22       if not identified :
23        sentence_tense = check_tense(sentence)
24        if sentence_tense == tf_nar_element[2]:
25         add sentence to tf_nar_element[1]
26        else:
27         add tf_nar_element to tfs_narrative
28         id=id+1
29         sentence_tense = check_tense(sentence)
30         tf_nar_element = ["tf_nar_"+id,[sentence],[sentence_tense]]
31         tocheck = 0
32      add tf_nar_element to tfs_narrative
33      return tfs_narrative

```

Figure 5. Narrative Timeframe Extraction.

Figure 5 provides the algorithm used for the Narrative Timeframes extraction. It takes as input the text, the patterns that identify the existence of a new timeframe, and the patterns that check the tense of a verb. The patterns that check the tense of the verbs are based on the part-of-speech tagging, dependencies, and the lemmatization of the verb. The lemmatization is the original form of a word without conjugation. We use it only to detect the verbs ‘be’ and ‘have’. Those elements are provided by Natural Language Processing tools such as Spacy [21]. For the part-of-speech tags of interest, we used:

- “VB” is assigned to the verbs base form
- “VBD” is assigned to verbs in the past tense
- “VBG” is the gerund (a verb that ends with ‘ing’)
- “VBN” assigned to the verb in past participle form
- “VBP” is assigned to the verbs in non-third person singular present form
- “VBZ” is assigned to the verbs in the third person singular present form

We considered the 12 principal tenses, and Table 2 provides some of the tenses and their respective patterns. We grouped the 12 verb tenses in the respective 5 tense categories: past anterior, past, present, future, future anterior [23]. For example, present continuous and present simple will both be present while present perfect, past simple, and past continuous will be considered as past. Based on the verb tense, the function check_tense will return the category of the verb tense identified.

TABLE II. SOME OF THE PATTERN USED TO DISTINGUISH VERB TENSE

| Verb Tense | tense_patterns |
|--------------------|--|
| Present Simple | pos = "VBZ" or pos = "VBP" |
| Present Continuous | verb with pos="VBG" and has_child = {dep= "AUX", pos = "VBZ" or "VBP", lemma="be"} |
| Past Simple | verb with pos="VBD" |
| Past Continuous | verb with pos="VBG" and has_child = {dep= "AUX", pos = "VBD", lemma="be"} |
| Future Simple | verb with pos="VB" and has_child = {dep= "AUX", pos = "MD", lemma="be"} |

As for the patterns that identify the presence of a new Narrative Timeframe, Table 3 presents some of them.

TABLE III. SOME OF THE PATTERNS THAT IDENTIFY NARRATIVE TIMEFRAMES

| Narrative_Patterns |
|--|
| A few ['years', 'months', 'days'] ['later', 'ago', 'back'] |
| ['earlier', 'later'] ['this', 'that'] ['years', 'months', 'days'] |
| In <date> |
| ['starting', 'from', 'starting from'] <date> |
| <number> ['years', 'months', 'days'] ['later', 'ago', 'back'] |

The algorithm goes as follows: an empty Narrative Timeframe is initialized. We go through all the sentences and we check the presence of a pattern. If no pattern is identified, we add the sentence to the timeframe. If patterns that trigger the creation of a new Narrative Timeframe are identified, we generate an identification to the new timeframe and we add to the previous timeframe to keep track of their connection. We then check the tense of the previous timeframe and the tense of the new one and we save the current timeframe element in the list of Narrative Timeframes. If the tenses are similar, we just add the sentence to the new timeframe. If the tenses are similar, there is a high risk that the author switches back to the previous timeframe. In that case, for the upcoming sentences, we keep track of any changes in the tenses. This is the only case in which the change of tense will trigger a change in the Narrative Timeframe. In future work, just like the event ordering approaches, a change in tense will trigger relations between events. Examples (12) and (13) clarify the need for this process:

- 12) *John is thinking about his life in New York. A few years ago, he had to move out because of his parents' job. He misses his friends dearly.*
- 13) *Alice graduated with a master's degree. A few months later, she found a job in an international company. She was finally able to move out.*

In 12), the change of the tense use can simulate a go back to the previous timeframe or just a need to change timeframes. In our current algorithm, we will just separate the three timeframes and handle the relationships between different timeframes in future works. As for 13), the

continuity in the tense simulates just a skip in time with no need for further tense monitoring of verb tenses. In the next section, we will be evaluating our approach.

V. EXAMPLE OF APPLICATIONS

In this section, we will present an output for each of the three algorithms provided above. Please note that for evaluating the output of the timeframe approach we used two types of resources, posts from LinkedIn and news scraped from multiple online news sites. For the LinkedIn posts, we focused on companies' accounts. We distinguish two types of posts, the short ones, and the long ones. For the long post, we selected 20 posts that contained multiple Narrative Timeframes or the need for Spoken Timeframes. As for news data, the scraping was made from three different sites in three different fields: politics, industry, and global. We took 10 news from each online source. The websites used were CNN (edition.cnn.com), BBC (bbc.com), and GlobalNews (globalnews.ca). We used this small amount of data in order to compare the output with the expected output manually since the approach is still in its early stages.

A. Publication Timeframe Extraction Results

The publishing timeframe extraction was not needed since the data was extracted "structured", with a distinction between the text and the date of publication. The Publication Timeframe was only applied to news data. For each site, we started by testing the performance of the Publication Timeframe since the scraper used is not customized for each website. We were able to identify the Publication Timeframe. Please note that a cleaning phase is necessary before applying the approach. Figure 6 presents one of the outputs of the algorithms. The figure provides part of the extracted text from a news site with the sentence with the pattern of interest highlighted in the input of the algorithm. We can notice the Publication Timeframe along with the rest of the text from which we removed the sentence with the pattern.



Figure 6. Output of Publication Timeframe Extraction Algorithm.

B. Spoken Timeframe Extraction Results

The Spoken Timeframe Extraction was applied to both data sets, the LinkedIn data, and the news data. Figure 7 provides the output of one of the provided data. Please note

that for better visualization, long paragraphs with no pattern were replaced by ‘...’ in the figure.

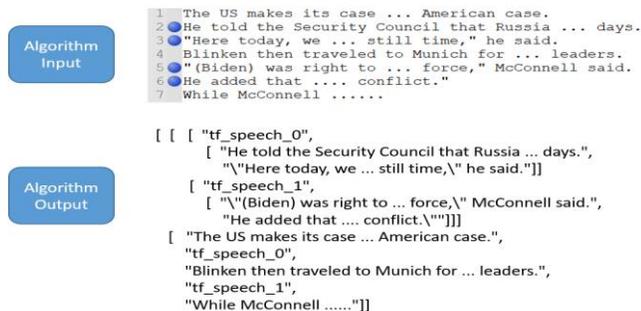


Figure 7. Output of Spoken Timeframe Extraction Algorithm.

In the example provided in Figure 7, we identified 4 sentences with patterns. They were distributed into 2 groups of consecutive sentences with patterns. The sentences in the example input were marked by blue dots next to them. In the output, we can notice that the algorithm provided a list with two elements in it, a Spoken Timeframe list and a list of the remaining sentences. The Spoken Timeframe list had 2 Spoken Timeframes in it, each having identification and consecutive sentences with patterns. As for the remaining text list, we notice that the extracted sentences were indeed replaced by their respective timeframe identification.

C. Narrative Timeframe Extraction Results

Finally, for the Narrative Timeframe, we used a text that had 2 Spoken Timeframes already identified in it. We also added a blue dot next to the sentences with the patterns identified. We can notice in the output the Narrative Timeframe list returned by our algorithm in Figure 8. It contains the three expected timeframes having their respective identification, the sentences ordered that belong to the timeframe, and the tense of the last sentence to enable comparisons.



Figure 8. Output of Spoken Timeframe Extraction Algorithm.

D. Results Analysis

Out of the 20 LinkedIn posts, the Spoken Timeframe was only available in one post due to the nature of posts on social media. But the Narrative Timeframe was performant, and we were able to segment the text as needed. We noticed one case of none-identified timeframes due to an unavailable pattern that was added later on to the model. As for the news data,

the Spoken Timeframes were highly identified in all fields due to the announcement or relay of speech of people. Out of the 30 news, only eleven showed multiple Narrative Timeframes. These timeframes were mostly used in news data when a follow-up on a story happens. In most cases, recent events are stated, before mentioning what has happened ‘earlier’ regarding the same story. For events that are anticipated to happen, the timeframes were not separated and will be considered in the future as the relation between events and not timeframes.

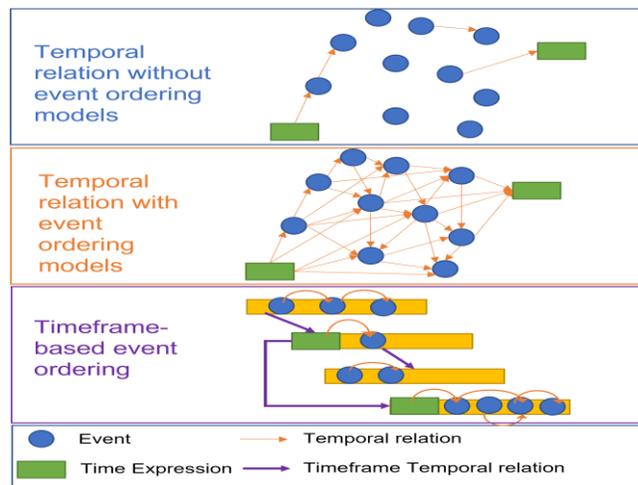


Figure 9. Temporal relations representation of multiple approaches.

Figure 9 shows the difference in the representation of temporal relations extract using multiple approaches. The first output shows a representation of an output without using an event ordering model. We can notice that hardly 4 events are connected and only 2 events are related to a time expression. The second presents the output of the same event extraction but after applying an event ordering model. This time most events are connected but the interpretation and the usability are complex. Finally, the third provides our desired representation using the temporal timeframes. This approach grouped events that occurred in the same period of time and limited the relation extraction between events from different timeframes. In this article we proposed the different type of timeframes and how to extract them without extracting the relationships between the timeframes.

VI. CONCLUSION

Finally, event extraction is an essential task in the NLP field. It enables the use of text data in order to build decision-making systems and for event monitoring. In this paper, we highlighted the need for timeframes to improve event ordering in the event extraction field. Three types of timeframes were presented: the publication, the narrative, and the Spoken Timeframe. Publication Timeframes will be used for multiple text analysis as a temporal indicator of the text. Narrative Timeframes enable the distinguishing of multiple periods of time used in a text, notably when a flashback or a flash-forward occurs. Finally, the Spoken Timeframe enables the distinction between the Narrative

Timeframes and the timeframe of “spoken” elements in a text, such as announcements or dialogs. We set a few patterns for the identification and extraction of the different timeframes. In future work, we will provide relation extraction methods for the timeframes and the events of each timeframe. We intend to evaluate the performance on longer texts and a larger number. We will also be distinguishing the multiple classes of events: point, process, culmination, and culminated process in order to identify states available in timeframes. This work will complete our study on detection and representation context from text [14].

REFERENCES

- [1] S. Bethard, “ClearTK-TimeML: A minimalist approach to TempEval 2013,” in *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, Atlanta, Georgia, USA, Jun. 2013, pp. 10–14.
- [2] S. W. Brown, J. Bonn, J. Gung, A. Zaenen, J. Pustejovsky, and M. Palmer, “VerbNet Representations: Subevent Semantics for Transfer Verbs”, in *Proceedings of the First International Workshop on Designing Meaning Representations*, Florence, Italy, Aug. 2019, pp. 154–163.
- [3] M. J. Buehner and J. May, “Knowledge mediates the timeframe of covariation assessment in human causal induction,” *Thinking & Reasoning*, vol. 8, no. 4, pp. 269–295, Nov. 2002, doi: 10.1080/13546780244000060.
- [4] N. Chambers, “NavyTime: Event and Time Ordering from Raw Text,” in *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, Atlanta, Georgia, USA, Jun. 2013, pp. 73–77. Accessed: Feb. 07, 2022. [Online]. Available: <https://aclanthology.org/S13-2012>
- [5] X. Duan *et al.*, “Temporality-enhanced knowledgememory network for factoid question answering,” *Frontiers Inf Technol Electronic Eng*, vol. 19, no. 1, Jan. 2018, pp. 104–115.
- [6] J. Hamann and L. Suckert, “Temporality in Discourse: Methodological Challenges and a Suggestion for a Quantified Qualitative Approach,” *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. Vol 19, p. No 2 (2018), Mar. 2018, doi: 10.17169/FQS-19.2.2954.
- [7] R. Han, Q. Ning, and N. Peng, “Joint Event and Temporal Relation Extraction with Shared Representations and Structured Prediction,” *arXiv:1909.05360 [cs]*, Sep. 2020.
- [8] F. Hogenboom, F. Frasincar, U. Kaymak, F. de Jong, and E. Caron, “A Survey of event extraction methods from text for decision support systems,” *Decision Support Systems*, vol. 85, pp. 12–22, May 2016, doi: 10.1016/j.dss.2016.02.006.
- [9] T. Kwiatkowski *et al.*, “Natural Questions: A Benchmark for Question Answering Research,” *Transactions of the Association for Computational Linguistics*, vol. 7, pp. 453–466, Aug. 2019, doi: 10.1162/tacl_a_00276.
- [10] N. Laokulrat, M. Miwa, Y. Tsuruoka, and T. Chikayama, “UTTime: Temporal Relation Classification using Deep Syntactic Features,” in *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, Atlanta, Georgia, USA, Jun. 2013, pp. 88–92.
- [11] A. Leeuwenberg and M.-F. Moens, “Structured Learning for Temporal Relation Extraction from Clinical Records,” in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, Valencia, Spain, Apr. 2017, pp. 1150–1158. Accessed: Feb. 09, 2022. [Online]. Available: <https://aclanthology.org/E17-1108>
- [12] Q. Li *et al.*, “Reinforcement Learning-Based Dialogue Guided Event Extraction to Exploit Argument Relations,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 520–533, 2022.
- [13] Y. Lu *et al.*, “Text2Event: Controllable Sequence-to-Structure Generation for End-to-end Event Extraction,” *arXiv:2106.09232 [cs]*, Jun. 2021.
- [14] No. Matta, N. Matta, E. Giret, and N. Declercq, “Enhancing Textual Knowledge Discovery using a Context-Awareness Approach”, in *The 8th Annual Conf. on Computational Science & Computational Intelligence (CSCI'21)*, December 15-17, 2021, USA, in press.
- [15] B. McDowell, N. Chambers, A. Ororbia II, and D. Reitter, “Event Ordering with a Generalized Model for Sieve Prediction Ranking,” in *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, Taipei, Taiwan, Nov. 2017, pp. 843–853.
- [16] Y. Meng and A. Rumshisky, “Context-Aware Neural Model for Temporal Information Extraction,” in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Melbourne, Australia, Jul. 2018, pp. 527–536. doi: 10.18653/v1/P18-1049.
- [17] M. Moens and M. Steedman, “Temporal Ontology and Temporal Reference,” *Computational Linguistics*, vol. 14, no. 2, pp. 15–28, 1988.
- [18] V. Moutinho, João Cordeiro, and P. Brazdil, “Association and Temporality between News and Tweets,” pp. 500–507, Sep. 2019, doi: 10.5220/0008362105000507.
- [19] W. H. Newton-Smith, “The Structure of Time,” *Routledge & CRC Press*, Jun. 01, 2020.
- [20] A. Pal, J. Margatan, and J. Konstan, “Question temporality: identification and uses,” in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, New York, NY, USA, Feb. 2012, pp. 257–260. doi: 10.1145/2145204.2145246.
- [21] E. Partalidou, E. Spyromitros-Xioufis, S. Doropoulos, S. Vologiannidis, and K. I. Diamantaras, “Design and implementation of an open source Greek POS Tagger and Entity Recognizer using spaCy”, in *2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Oct. 2019, pp. 337–341.
- [22] T. Poell, “Social media, temporality, and the legitimacy of protest,” *Social Movement Studies*, vol. 19, no. 5–6, pp. 609–624, Nov. 2020, doi: 10.1080/14742837.2019.1605287.
- [23] Z. Vendler, “Verbs and Times,” *The Philosophical Review*, vol. 66, no. 2, pp. 143–160, 1957, doi: 10.2307/2182371.
- [24] Z. Vendler, *Linguistics in Philosophy*. Cornell University Press, 2019. doi: 10.7591/9781501743726.
- [25] X. Wu, T. Wang, Y. Fan, and F. Yu, “Chinese Event Extraction via Graph Attention Network,” *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, vol. 21, no. 4, p. 71:1-71:12, Jan. 2022, doi: 10.1145/3494533.
- [26] W. Xiang and B. Wang, “A Survey of Event Extraction From Text,” *IEEE Access*, vol. 7, pp. 173111–173137, 2019
- [27] J. Zagal and M. Mateas, “Time in Video Games: A Survey and Analysis,” *Simulation & Gaming - Simulat Gaming*, vol. 41, Jan. 2010, doi: 10.1177/1046878110375594.
- [28] W. Zhao, J. Zhang, J. Yang, T. He, H. Ma, and Z. Li, “A novel joint biomedical event extraction framework via two-level modeling of documents,” *Information Sciences*, vol. 550, pp. 27–40, Mar. 2021, doi: 10.1016/j.ins.2020.10.047.

Linear Fuzzy Space Based Framework for Air Quality Assessment

Endre Pap, Đorđe Obradović, Zora Konjović, Ivan Radosavljević

Singidunum University
Belgrade, Serbia

Emails: {epap, djobradovic, zkonjovic, iradosavljevic}@singidunum.ac.rs

Abstract— Air quality is one of the most critical issues humankind is facing today. There are diverse types of indices measuring the air pollution which are mainly based on aggregation functions. This paper proposes a model aimed at forecasting aggregated air pollution indices based on our theory of the linear fuzzy space. The proposed original model consists essentially of two sub models. The first one models concentrations of pollutants, while the second one models Air Quality Index (AQI). We model concentrations of pollutants by regression (XGBoost and deep neural network) utilizing fuzzy time series of two groups of data (measured concentrations and meteorological parameters). Multi-contaminant air quality index is modelled as an aggregation of Pollutant Standard Index (PSI) obtained via fuzzy linear transformation defined by fuzzy breakpoints. Some preliminary results are presented indicating model performance in terms of prediction mean absolute errors.

Keywords: *fuzzy set; linear fuzzy space; AQI index; aggregation operator.*

I. INTRODUCTION

In the last decade, mankind has been facing air pollution as one of the most important issues with adverse effect on human health, but also on the economy of societies. According to WHO's annual World Health Statistics report from 2016, outdoor air pollution causes approximately 4.2 million deaths per year [1]. As reported by the European Environmental Agency (EEA) in 2018, the number of deaths in Europe related to concentrations of the particles PM2.5 was about 379,000 [2]. Therefore, there is a great need for the air pollution forecasting models which will express the air pollution as a simple value that is understandable for a wide audience.

Air pollution is an extremely complex spatio-temporally determined dynamic system distinctly characterized by the presence of imprecision and uncertainty. Therefore, it is not easy to give a precise air pollution forecast, which would be of great importance for public health.

To cope with uncertainty and imprecision, we use a fuzzy approach. More precisely, the one based on our previous results presented in [3] - [8], where we use mathematical models for basic concepts: fuzzy point, fuzzy spatial relation, fuzzy ordering, and fuzzy distance. For modelling the temporal dimension of air pollution, we use a combination of time series models with techniques supporting the manipulation of imprecise and uncertain data, known under the umbrella term Fuzzy Time Series (FTS). This model enables a more adequate air pollution forecast.

Multi-contaminant Air Quality Index (AQI) manages multiple effects due to the exposure to more pollutants, gives more complete information on the possible impacts of air pollutants and a direction for a more accurate, consistent, and comparable AQI system. Hence, we opt for multi-contaminant AQIs as a model of air pollution estimate.

The rest of the paper is organized in four sections. Section 2 presents theoretical foundations together with related work, while Section 3 presents the model of the proposed framework. Section 4 presents the simulation results for the real data set (82457 samples/16 variables/24h measurements) while Section 5 summarizes the research results, identifies model deficiencies, and outlines future research.

II. PRELIMINARIES

The theoretical foundations of our model are based on linear fuzzy space theory, multi-contaminant fuzzy AQIs, and fuzzy time series.

A. Linear fuzzy space

In this subsection, we present the fundamental concepts of the linear fuzzy space: fuzzy point, linear fuzzy space, fuzzy space ordering, fuzzy space metrics, and fuzzy linear combination, as defined in [3] – [7].

Definition 1 Fuzzy point $P \in R^2$, denoted by \tilde{P} is defined by its membership function $\mu_{\tilde{P}} \in \mathcal{F}^2$, where the set \mathcal{F}^2 contains all membership functions $u: R^2 \rightarrow [0,1]$ satisfying the following conditions:

- i) $(\forall \mu \in \mathcal{F}^2)(\exists_1 P \in R^2) \mu(P) = 1$,
- ii) $(\forall X_1, X_2 \in R^2)(\lambda \in [0,1]) \mu(\lambda X_1 + (1 - \lambda)X_2) \geq \min(\mu(X_1), \mu(X_2))$,
- iii) function μ is upper semi-continuous,
- iv) $[\mu]^\alpha = \{X | X \in R^2, \mu(X) \geq \alpha\}$ α -cut of function μ is convex.

Here, a point from R^2 with a membership function $\mu_{\tilde{P}}(P) = 1$, is denoted by P (P is the core of the fuzzy point \tilde{P}), the membership function of point \tilde{P} is denoted by $\mu_{\tilde{P}}$, while $[P]^\alpha$ stands for the α -cut (a set from R^2) of the fuzzy point.

Definition 2 R^2 Linear fuzzy space is the set $\mathcal{H}^2 \subset \mathcal{F}^2$ of all functions which, in addition to the properties given in Definition 1, are:

- i) Symmetrical with respect to the core $S \in R^2$ ($\mu(S) = 1$),
 $\mu(V) = \mu(M) \wedge \mu(M) \neq 0 \Rightarrow d(S, V) = d(S, M)$,
where $d(S, M)$ is the distance in R^2 .

- ii) Inverse-linearly decreasing regarding points' distance from the core, i.e.:

$$\text{If } r \neq 0: \mu_S(V) = \max\left(0, 1 - \frac{d(S,V)}{|r_S|}\right),$$

$$\text{If } r = 0: \mu_S(V) = \begin{cases} 1 & \text{if } S = V \\ 0 & \text{if } S \neq V, \end{cases}$$

where $d(S, V)$ is the distance between point V and the core S ($V, S \in R^2$) and $r \in R$ is a constant.

The elements of that space are represented as ordered pairs $\tilde{S} = (S, r_S)$ where $S \in R^2$ is the core of \tilde{S} , and $r_S \in R$ is the distance from the core for which the function value becomes 0.

Measurement in the space, especially the distance between plane geometry objects, is defined as a generalization of the concept of physical distance:

Definition 3 Let \mathcal{H}^2 be a linear fuzzy space and $\tilde{d}: \mathcal{H}^2 \times \mathcal{H}^2 \rightarrow \mathcal{H}^+$, $L, R: [0,1] \times [0,1] \rightarrow [0,1]$ be symmetric, associative, and non-decreasing for both arguments, and $L(0,0) = 0$, $R(1,1) = 1$. The ordered quadruple $(\mathcal{H}^2, \tilde{d}, L, R)$ is called fuzzy metric space and the function \tilde{d} is a *fuzzy metric*, if and only if the following conditions hold:

- i) $\tilde{d}(\tilde{X}, \tilde{Y}) = \tilde{d}(\tilde{Y}, \tilde{X}) \Leftrightarrow [\tilde{X}]^1 = [\tilde{Y}]^1$
- ii) $\tilde{d}(\tilde{X}, \tilde{Y}) = \tilde{d}(\tilde{Y}, \tilde{X})$, $\forall \tilde{X}, \tilde{Y} \in \mathcal{H}^2$
- iii) $\forall \tilde{X}, \tilde{Y} \in \mathcal{H}^2$:

$$\tilde{d}(\tilde{X}, \tilde{Y})(s+t) \geq L(d(x,z)(s), d(z,y)(t))$$

$$\text{if } s \leq \lambda_1(x,z) \wedge t \leq \lambda_1(z,y) \wedge s+t \leq \lambda_1(x,y)$$

$$\tilde{d}(\tilde{X}, \tilde{Y})(s+t) \leq R(d(x,z)(s), d(z,y)(t))$$

$$\text{if } s \geq \lambda_1(x,z) \wedge t \geq \lambda_1(z,y) \wedge s+t \geq \lambda_1(x,y),$$

The α -cut of a fuzzy number $\tilde{d}(x, y)$ is given by

$$[\tilde{d}(\tilde{X}, \tilde{Y})]^\alpha = [\lambda_\alpha(x, y), \rho_\alpha(x, y)] \quad (x, y \in R^+, 0 < \alpha \leq 1).$$

The fuzzy zero, $\tilde{0}$ is a non-negative fuzzy number with $[\tilde{0}]^1 = 0$.

Definition 4 Let \mathcal{H}^2 be a linear fuzzy space. Then, function $f: \mathcal{H}^2 \times \mathcal{H}^2 \times [0,1] \rightarrow \mathcal{H}^2$ called a *linear combination* of the fuzzy points $\tilde{A}, \tilde{B} \in \mathcal{H}^2$ is given by:

$$f(\tilde{A}, \tilde{B}, u) = \tilde{A} + u \cdot (\tilde{B} - \tilde{A}),$$

where $u \in [0,1]$ and the operator \cdot is the scalar multiplication of the fuzzy point.

B. Fuzzy Air pollution indices

As shown in [7], a multi-contaminant model of AQI, in which aggregation functions (aggregation operators) are applied to combine several numerical values into a single representative, is predominant by far. An aggregation operator has natural properties such as monotonicity and boundary conditions. In practice, the data is usually normalized, so the definition of aggregation becomes:

Definition 5. An aggregation function (operator) is a function $A^{(n)}: [0,1]^n \rightarrow [0,1]$ which satisfies the following conditions:

1. is nondecreasing (in each variable)
2. $A^{(n)}(0, \dots, 0) = 0$ and $A^{(n)}(1, \dots, 1) = 1$.

Aggregation applies to various fields and takes diverse forms, from the simple to quite sophisticated ones, modelling the interaction between criteria which are managed by

monotone set functions and corresponding integrals [8], [10-12] that will be incorporated in our future research.

The simplest *AQI* model calculates a sub-index (AQI_i) for each pollutant i by the following linear interpolation formula:

$$AQI_i = \frac{I_{high} - I_{low}}{C_{high} - C_{low}}(C - C_{low}) + I_{low}.$$

Here, C is the monitored ambient average concentration of pollutant i ; C_{low} is the breakpoint lower than or equal to C ; C_{high} is the breakpoint higher than or equal to C ; and I_{low} and I_{high} are the sub-index values corresponding to C_{low} and C_{high} , respectively. The overall *AQI* is then calculated as a simple max aggregation:

$$AQI = \max_{i=1}^m (AQI_i).$$

There is ongoing research for new aggregation functions, which involve the influence of multiple pollutants (see [13] - [20]). Among these *AQIs*, arithmetic pollutant aggregation integrates pollutants in a linear or nonlinear way, and weighted pollutant aggregation further assigns varied weights from different approaches. The General Air Quality Health Index (*GAQHI*) is proposed as a pollutant-aggregated, local health-based *AQI* paradigm suitable for representing a complex multi-contaminant situation:

$$I_s = \left(\sum_{i=1}^n (AQI_i)^\alpha \right)^{\frac{1}{\alpha}},$$

where $\alpha \in [1, \infty]$.

An interesting modification of the United States Environmental Protection Agency (EPA) *AQI* is proposed in [13], giving a new index *RAQI* which is the product of three terms:

$$RAQI = F_1 * F_2 * F_3$$

were

$$F_1 = \max(I_i), \quad i = 1, 5$$

$$F_2 = \frac{\sum_{i=1}^5 Ave_{daily}(I_i)}{Ave_{annual} \cdot \left(\sum_{i=1}^5 Ave_{daily}(I_i) \right)}$$

and the Shannon entropy function is introduced in the third term:

$$F_3 = \frac{Ave_{annual} \cdot Entropy_{daily}(\max_{i=1}^5(I_i))}{Entropy_{daily}(\max_{i=1}^5(I_i))}$$

This model strives to avoid ambiguity (indicating a less polluted air as highly polluted) and ellipticity (indicating highly polluted air as less polluted) by introducing entropy.

In addition, there are interesting approaches like in [14] that model the air pollution index via a mixture of distributions based on its structure and descriptive status. Each of these models can be easily fuzzified by simply mapping crisp space to a linear fuzzy space by the suitable fuzzy function over a crisp domain.

There are also results that utilize fuzzy logic for modelling air quality indices, like those presented in [15], [16], [19], and [20].

In [15], the input variables are air pollutant criteria (PM10, SO2, CO, NO2, O3), and the output variable is fuzzy AQI.

The fuzzification process is defined via the boundary values of the universal sets and the corresponding fuzzy sets (trapezoidal for input, and triangular for output variables). The rule base representing the relationship between input variables and output variables contains 243 rules. The max-min inference strategy and centroid method are chosen for the inference and defuzzification process.

In [16], ten parameters (selected concentrations of pollutants) are divided into two groups. Firstly, the parameters in each group are processed by the inference systems, and then grouped and normalized between 0 and 100, resulting in two new groups. These new groups are processed in the second step by new inference systems and result in the Fuzzy Air Quality Index (FAQI). All rules (72 in total) have only one antecedent. A classical fuzzy inference system (Mamdani) is used for this purpose. Common to those results is a reliance on an approach that utilizes fuzzy rule-based inference and does not include a time series-based prediction.

In [19], the authors present a comparative study of the results obtained from several models for air pollution index forecast which shows that the fuzzy time series models outperformed the other models in terms of forecasting accuracy and computation time. Finally, [20] utilizes a fuzzy time series-Markov chain model for predicting the daily air pollution index.

C. *Fuzzy time series*

Most of the real-world tasks that utilize time series rely on multivariate time series models [21]-[24]. The common multivariate time series model is [21]:

Let $Z_t = [Z_{1,t}, Z_{2,t}, \dots, Z_{m,t}]$ be an m -dimensional jointly stationary real-valued vector process such that $E(Z_{i,t}) = \mu_i$ is a constant for each $i = 1, 2, \dots, m$ and the cross-covariances between $Z_{i,t}$ and $Z_{j,s}$ for all $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, m$ are functions only of the time difference $(s - t)$.

On the other hand, the original definition of the univariate first fuzzy time series model is [24]:

Definition 6. Let $Y(t) (t = \dots, 0, 1, 2, \dots)$, a subset of R^1 be the universe of discourse on which fuzzy sets $f_i(t) (i = 1, 2, \dots)$ are defined and $F(t)$ is the collection of $f_i(t) (i = 1, 2, \dots)$. Then, $F(t)$ is called a fuzzy time series on $Y(t) (t = \dots, 0, 1, 2, \dots)$.

Our time series model is a combination of the previous two where we apply the same common multivariate model which is modified to support imprecise values. In our model, we simply replace a crisp point with a linear fuzzy space point [2]:

Definition 7. Let $Y(t) (t = \dots, 0, 1, 2, \dots)$, a subset of R^1 be the universe of discourse. Let $H^l (l = 1, 2)$ be a linear fuzzy space. Furthermore, let $f_i(t) (i = 1, 2, \dots)$ be fuzzy sets defined as points on a linear fuzzy space over the given universe of discourse, and $\tilde{F}_j(t) (j = 1, 2, \dots, m)$ be collections of these fuzzy points. Then, $\tilde{F}_t = [\tilde{F}_{1,t}, \tilde{F}_{2,t}, \dots, \tilde{F}_{m,t}]'$ is called a linear fuzzy space based fuzzy time series on $Y(t) (t = \dots, 0, 1, 2, \dots)$. This definition enables all features of linear fuzzy space to be utilized. For example, a process vector can

be of a mixed type (some components can be crisp, some can be fuzzy) whilst spatial relations defined on the linear fuzzy space hold.

In our framework, diverse machine learning techniques can be used to create complex, non-linear relations.

III. FUZZY MODEL OF AIR POLLUTION INDICES PREDICTION

In this example, we demonstrate how the linear fuzzy space is used for time series-based forecasting. Fuzzy time series defined by means of the fuzzy linear space, as described in subsection C, are used to model air quality forecast.

A. *Data model*

The data model used in this paper consists of temporal georeferenced samples. Each sample is a time series covering the previous 24h in 1h sample rate (total 385 real values). Each time series corresponds to one variable. Variables are divided in two groups: meteorological parameters, the Global Data Assimilation System (GDAS), and six common air pollutants known as “criteria air pollutants.”

GDAS parameters [25] are described in Table I.

TABLE I. GDAS PARAMETERS

| ID | Description | Unit |
|------------|--|------|
| PRSS | Pressure at surface | hPa |
| TPP6 | Accumulated precipitation (6 h accumulation) | m |
| RH2M | Relative Humidity at 2m AGL | % |
| TO2M | Temperature at 2m AGL | K |
| TCLD | Total cloud cover (3- or 6-h average) | % |
| U10M | U-component of wind at 10 m AGL | m/s |
| V10M | V-component of wind at 10 m AGL | m/s |
| TMPS | Temperature at surface | K |
| PBLH | Planetary boundary layer height | m |
| irradiance | Irradiance/solar power | W/m2 |

Criteria air pollutants [26] are given in Table II.

TABLE II. AIR POLLUTANTS PARAMETERS

| ID | Description | Unit |
|------|--|--------------------------|
| PM10 | Suspended particles smaller than 10 μm | $\mu\text{g}/\text{m}^3$ |
| PM25 | Suspended particles smaller than 2.5 μm | $\mu\text{g}/\text{m}^3$ |
| SO2 | Sulphur dioxide | ppb |
| CO | Carbon Monoxide | ppm |
| NO2 | Nitrogen Dioxide | ppb |
| O3 | Ground-level Ozone | ppm |

B. *Linear fuzzy space-based air pollution index*

Since air pollutants are measured in different physical units and scales, the first step is to transform them into a common domain (0-500). This transformation is usually defined by breakpoint tables and the resulting values are called Pollutant Standard Index (PSI). Instead of using discrete functions, we propose a fuzzy linear transformation defined by fuzzy breakpoints (Figure 1).

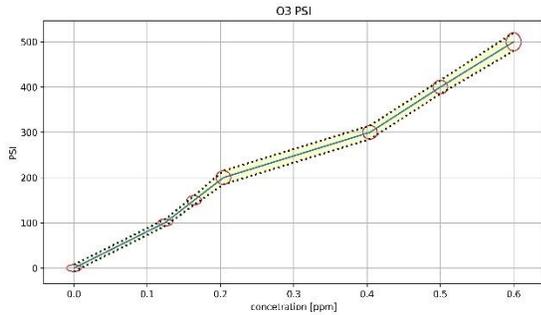


Figure 1. Fuzzy linear transformation

A fuzzy linear transformation is defined by an ordered list of 2D Fuzzy points $\tilde{P} = (\tilde{X}, \tilde{Y})$. Each 2D fuzzy point consists of two components $\tilde{X} = (X, r_x)$ and $\tilde{Y} = (Y, r_y)$ which are 1D fuzzy points. Then, Fuzzy PSI (FPSI) is defined as:

$$\widetilde{FPSI}_i = \widetilde{interp}(C, [\tilde{P}_0, \dots, \tilde{P}_n]) = (FPSI_i, r_{FPSI}_i),$$

$$FPSI_i = \frac{Y_{high} - Y_{low}}{X_{high} - X_{low}} (C - X_{low}) + Y_{low}$$

$$r_{PSI} = \frac{r_{Yhigh} - r_{Ylow}}{r_{Xhigh} - r_{Xlow}} (C - X_{low}) + r_{Ylow}$$

where \widetilde{interp} is a fuzzy linear transformation from concentration fuzzy space into index fuzzy space. Fuzzy points \tilde{P}_{high} and \tilde{P}_{low} are fuzzy points whose roots of \tilde{X} components are nearest to the concentration C .

$FPSI$ can be further represented by a linguistic variable, or it can be used directly in the aggregation process.

A single fuzzy value $FAQI$ is obtained by applying some fuzzy aggregation operator (aggreg) to all (n) component $FPSI$ indices:

$$FAQI = \text{aggreg}(FPSI_i), i = 1, n$$

To simplify the decision-making process and/or facilitate general understanding, a fuzzy linguistic variable defined by corresponding fuzzy sets can be easily introduced in such a model.

C. Prediction model

In our model, we opt for multivariate regression to forecast $FAQI$ (Figure 2). However, other classification methods can easily be incorporated in the proposed model.

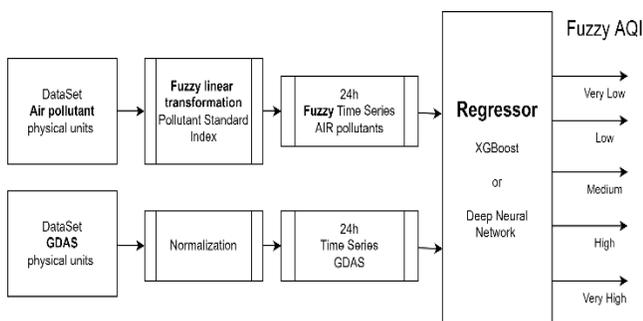


Figure 2. Prediction model

IV. MODEL APPLICATION

In order to present the proposed model/methodology, we ran one experiment on a large and diverse data set. The used data set contains more than 82000 samples each with 385 real values. GDAS values are interpolated to fit five geo locations and merged with measurements of the concentration of the air pollutants.

A. Data set

In this experiment, we used five data sets from five distinct locations in USA, each in the same format. The sources of data are [25] and [26]. Samples are indexed by temporal attribute, datetime, ranging from January 1, 2015 to December 31, 2021. All ten meteorological GDAS and six air pollutants are stored in 24 hours' time slot with 1h sample rate (385 real values in total). Table III presents the data in more details (sample sizes per locations).

TABLE III. DATA SETS

| site_id | site | Samples |
|-------------|------------------|---------|
| 11-001-0043 | Washington, DC | 27,981 |
| 13-089-0002 | Near Atlanta, GA | 21,468 |
| 18-097-0078 | Indianapolis, IN | 16,774 |
| 22-033-0009 | Baton Rouge, LA | 6,569 |
| 32-003-0540 | Las Vegas, NV | 9,665 |

The same source provides data about land use (COMMERCIAL, RESIDENTIAL) and type of location (URBAN, SUBURBAN, RURAL), as shown in Table IV.

TABLE IV. SITE TYPES

| ID | City | Land use | Location |
|-------------|------------------|-------------|----------|
| 11-001-0043 | Washington, DC | COMMERCIAL | URBAN |
| 13-089-0002 | Near Atlanta, GA | RESIDENTIAL | SUBURBAN |
| 18-097-0078 | Indianapolis, IN | RESIDENTIAL | SUBURBAN |
| 22-033-0009 | Baton Rouge, LA | COMMERCIAL | URBAN |
| 32-003-0540 | Las Vegas, NV | RESIDENTIAL | URBAN |

PSI calculation was done using PSI functions (Table V), which transform the physical value domain into a real value interval [0, 500].

TABLE V. PSI BREAKPOINTS

| PSI | PM10 $\mu\text{g}/\text{m}^3$ | SO2 ppm | CO ppm | NO2 ppm | O3 ppm |
|-----|----------------------------------|------------|-----------|------------|-----------|
| 0 | 0 | 0 | 0 | 0 | 0 |
| 50 | 50 | 0.03 | 4.5 | - | 0.06 |
| 100 | 150 | 0.14 | 9 | - | 0.12 |
| 200 | 350 | 0.3 | 15 | 0.6 | 0.2 |
| 300 | 420 | 0.6 | 30 | 1.2 | 0.4 |
| 400 | 500 | 0.8 | 40 | 1.6 | 0.5 |
| 500 | 600 | 1 | 50 | 2 | 0.6 |

B. Fuzzy air quality index

In our framework, the fuzzy air quality index is modelled via a simple max aggregation function applied to five $FPSI$ indices of each criteria air pollutants:

$$FAQI = \max(FPSI_{CO}, FPSI_{PM10}, FPSI_{NO2}, FPSI_{O3}, FPSI_{SO2})$$

Finally, we introduce a fuzzy linguistic variable (*very low, low, medium, high, very high*) defined by corresponding fuzzy sets, as depicted in Figure 3.

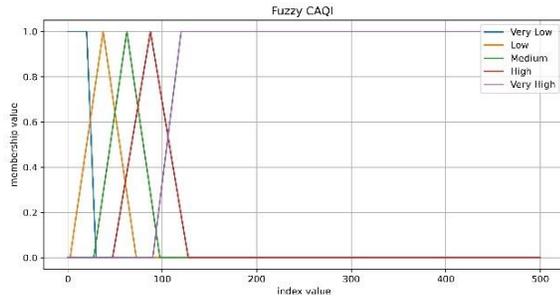


Figure 3. Fuzzy CAQI

C. Prediction

In our experiments, we applied two multivariate regressors. The first multivariate predictor regressor in this experiment is XGBoostRegressor with 24*10 GDAS and 24*6 air pollutant variables as input, and 5 real valued outputs, each corresponding to a single fuzzy set (FAQL *very low* to *very high*), as depicted in Figure 3. Data set is split up into train (80%) and test (20%) subsets and trained with 1000 estimators with *max_depth* 4 and enabled early stopping method to avoid overfitting.

The second multivariate predictor regressor in this experiment is a deep neural network with 24*10 GDAS and 24*6 air pollutant variables as input, and 5 real valued outputs, each corresponding to a single fuzzy set (FAQL *very low* to *very high*), with one hidden layer consisting of 20 Rectified Linear Units (ReLU) nodes. The activation functions in output layer are Sigmoid. The data set is split up into train (80%) and test (20%) subsets. Two dropout layers with 10% random filters are intercepted between active layers to prevent overfitting.

D. Prediction results

The mean absolute errors for FAQI prediction are shown in Table VI (XGBoost) and Table VII (deep neural network).

TABLE VI. XGBOOST

| ID | very low | low | medium | high | very high |
|-------------|----------|-------|--------|-------|-----------|
| 11-001-0043 | 0.229 | 0.230 | 0.049 | 0.003 | 0.001 |
| 13-089-0002 | 0.239 | 0.217 | 0.036 | 0.004 | 0.001 |
| 18-097-0078 | 0.213 | 0.213 | 0.065 | 0.006 | 0.001 |
| 22-033-0009 | 0.224 | 0.218 | 0.063 | 0.009 | 0.000 |
| 32-003-0540 | 0.083 | 0.236 | 0.185 | 0.045 | 0.012 |

TABLE VII. DEEP NEURAL NETWORK

| ID | very low | Low* | medium | high | very high |
|-------------|----------|-------|--------|-------|-----------|
| 11-001-0043 | 0.405 | 0.399 | 0.065 | 0.003 | 0.001 |
| 13-089-0002 | 0.460 | 0.398 | 0.045 | 0.004 | 0.002 |
| 18-097-0078 | 0.425 | 0.406 | 0.074 | 0.005 | 0.002 |
| 22-033-0009 | 0.426 | 0.392 | 0.056 | 0.008 | 0.001 |
| 32-003-0540 | 0.109 | 0.362 | 0.291 | 0.039 | 0.010 |

Tables VI and VII show that both regressors behave similarly. Moreover, they are good in prediction for categories *medium, high* and *very high* and poor in prediction for categories *very low* and *low*. Having that in mind and the main purpose of the FAQI to alarm of dangerous air pollution (*high* and *very high*, possibly *medium*), the results indicate that further research is needed and justified.

V. CONCLUSION

This paper proposes a model aimed at forecasting the aggregated air pollution index that is based on our theory of the linear fuzzy space. The proposed model consists of two sub models. The first one models the concentration of pollutants, while the second one models multi-contaminant air quality index. We model the concentration of pollutants by regression, utilizing fuzzy time series of two groups of data: measured concentrations of pollutants and meteorological parameters. The multi-contaminant air quality index is modeled as a fuzzy aggregation of PSI obtained via fuzzy linear transformation defined by fuzzy breakpoints.

Preliminary results show that our model is characterized by a distinct property which is a good performance for higher values of air quality index, and significantly worse (mean absolute errors higher for an order of magnitude) performance for lower values. This is a notable deficiency of the model calling for improvement that will ensure equally good performance for all categories.

Air pollution is a result of an extremely complex and interdependent interaction among multiple factors (air pollutants, environment, time, climate conditions, etc.) additionally burdened with uncertainty and imprecision in data. This makes a single index a rough approximation of the considered pollution situation.

Indeed, there is a potential for improvements in the research topics tackled in this paper which shapes further research directions. The possible improvements could be further divided in two rough partitions. The first, which is of fundamental kind, is about rethinking the air pollution index concept (for example, making it contextually dependent, or making it multidimensional). The second one is about improvement of the model proposed in this paper: use of new parameters (like those in Table IV), training data balancing, learning shapes of membership functions from historical data, and alike. Improvements should specifically address creation of precision metrics in linear fuzzy space, enabling estimations of sensitivity of interval partitions selection in time series, aggregation models, and fuzzy sets parameters. Recent theory development (see [27], [28]) gives a method for identification of the optimal solution for convex and non-convex optimization in fuzzy approach that could help to do this.

The two partitions intersect at utilization of Artificial Intelligence (AI) methods, particularly fuzzy approach, and machine learning techniques.

ACKNOWLEDGMENT

The authors acknowledge funding provided by the Science Fund of the Republic of Serbia #GRANT No. 6524105, AI - ATLAS.

REFERENCES

- [1] World health statistics 2016: monitoring health for the SDGs, sustainable development goals. 1.Health Status Indicators. 2.Global Health. 3.Health Priorities. 4.Mortality. 5.Universal Coverage. 6.Life Expectancy. 7.Statistics. I.World Health, E-ISBN 978 92 4 069569 6, <https://www.who.int/docs/default-source/gho-documents/world-health-statistic-reports/world-health-statistics-2016.pdf> [Accessed: May 15th 2022]
- [2] Air quality in Europe – 2018 report, ISSN 1977-8449, https://www.eea.europa.eu/publications/air-quality-in-europe-2018/at_download/file [Accessed May 15th, 2022]
- [3] Đ. Obradović, "Model of the computer simulation system for managing geospace under uncertainty conditions," PhD thesis, Faculty of Technical Sciences University of Novi Sad (in Serbian), 2011.
- [4] Đ. Obradović, Z. Konjović, E. Pap, and A. Šošarić, "The linear fuzzy space: Theory and applications," in E.Pap, Ed., *Artificial Intelligence: Theory and Applications*, Springer, 2021, pp 227-253,
- [5] Đ. Obradović, Z. Konjović, E. Pap, and I. Rudas. "Fuzzy Geometry in Linear Fuzzy Space, Intelligent Systems: Models and Applications," Revised and Selected Papers from the 9th IEEE International Symposium on Intelligent Systems and Informatics SISY 2011. Springer Berlin Heidelberg, 2013, pp 137-153.
- [6] Đ. Obradović, Z. Konjović, E. Pap, and N. Ralević "The maximal distance between imprecise point objects," *Fuzzy Sets and Systems* 170 (1), pp. 76-94, 2011, doi: <https://doi.org/10.1016/j.fss.2010.12.005>
- [7] Đ. Obradović, Z. Konjović, E. Pap, and I. J. Rudas "Linear Fuzzy Space Based Road Lane Detection" *Knowledge Based Systems* 38 (1), pp. 37-47, 2013, doi: <https://doi.org/10.1016/j.knosys.2012.01.002>
- [8] M. Grabisch, J. L. Marichal, R. Mesiar, and E. Pap, "Aggregation functions (Encyclopedia of mathematics and its applications, series number 127)," Cambridge University Press, 2009.
- [9] X. Tan et al., "A review of current quality indexes and improvements under the multi-contaminant air pollution exposure," *J. Environ. Manag.* 279, <https://doi.org/10.1016/j.jenvman.2020.111681> 2021.
- [10] E. P. Klement, R. Mesiar, and E. Pap, "Triangular norms," Kluwer, 2000.
- [11] E. Pap (Ed.), "Artificial Intelligence: Theory and Applications," Springer, 2021.
- [12] E. Pap (Ed.), "Intelligent Systems: Models and Applications. Topics in Intelligent Engineering and Informatics 3," Springer, 2013.
- [13] W. Cheng, "Revised air quality index derived from an entropy function," *Atmospheric Environment*, 38(3), pp 383-391, 2004
- [14] N. A. AL-Dhurafi, N. Masseran, Z. H. Zamzuri, and M. A. Mohd Safari "Modeling the Air Pollution Index based on its structure and descriptive status," *Air Qual Atmos Health* 11, pp 171–179, 2018. <https://doi.org/10.1007/s11869-017-0528-21>
- [15] I. Atacak, N. Arici, and D. Guner, "Modelling and evaluating air quality with fuzzy logic algorithm-Ankara-Cebeci sample," *Int J Intell Syst Appl Eng*, vol. 5, no. 4, pp. 263–268, Dec. 2017.
- [16] A. Fino , F. Vichi, Ch. Leonardi, and K. Mukhopadhyay, "An overview of experiences made and tools used to inform the public on ambient air quality," *Atmosphere* 12(1), pp 1: 1524, 2021, Available: <https://doi.org/10.3390/atmos12111524>, [Accessed: January 12, 2022].
- [17] G. Kyrkilis, A. Chaloulakou, and P.A. Kassomenos, "Development of an aggregate air quality index for an urban mediterranean agglomeration: Relation to potential health effects," *Environ. Int.* 33, pp 670–676, 2007.
- [18] M. H. Sowlat, H. Gharibi, M. Yunesian, M. T. Mahmoudi, and S. Lotfi "A novel, fuzzy-based air quality index (FAQI) for air quality assessment," *Atmospheric Environment*, 45(12), pp. 2050-2059, 2011, <https://doi.org/10.1016/j.atmosenv.2011.01.060>.
- [19] J.W. Koo, S.W. Wong, G. Selvachandran, H. V. Long, and L. H. Son, "Prediction of Air Pollution Index in Kuala Lumpur using fuzzy time series and statistical models," *Air Qual Atmos Health* 13, pp 77–88, 2020. <https://doi.org/10.1007/s11869-019-00772-y>
- [20] Y. Alyousifi, M. Othman, R. Sokkalingam, I. Faye, and P. C. L. Silva, "Predicting Daily Air Pollution Index Based on Fuzzy Time Series Markov Chain Model," *Symmetry* 2020, vol. 12, no. 293, 2020; doi:10.3390/sym12020293
- [21] W. W. S. Wei, "Multivariate Time Series Analysis and Applications", John Wiley & Sons Ltd., Book Series: Wiley Series in Probability and Statistics, 2019, doi:10.1002/9781119502951
- [22] Q. Song and B.S. Chissom, "Fuzzy time series and its model", *Fuzzy Sets and Systems* vol.3 no. 54, pp. 269–277, March 1993, doi:10.1016/0165-0114(93)90372-O
- [23] S. Panigrahi and H.S. Behera, "Fuzzy Time Series Forecasting: A Survey," in H. Behera, J. Nayak, B. Naik, and D. Pelusi Eds, *Computational Intelligence in Data Mining. Advances in Intelligent Systems and Computing*, vol 990. Springer, Singapore, 2020, https://doi.org/10.1007/978-981-13-8676-3_54
- [24] A. Sahin, M. F. Dodurka, T. Kumbasar, E. Yesil, and S. Siradag, "Review Study on Fuzzy Time Series and Their Applications in the Last Fifteen Years," *International Fuzzy Systems Symposium (FUZZYSS'15)*, Istanbul, 4, pp 166-171, 2015.
- [25] NOAA-Air Resources Laboratory, Global Data Assimilation System (GDAS1) [Data set], Archive began: December 1th 2004, <https://www.ready.noaa.gov/gdas1.php>, [Accessed: January 10, 2022]
- [26] EPA United States Environment Protection Agency, eGRID2020 [Data set], January 27th 2022, <https://www.epa.gov/egrid/download-data> [Accessed: February 1, 2022].
- [27] S. Das and D. Chakraborty, *An Introduction to Analytical Fuzzy Plane Geometry*, STUDEFUZZ, vol. 381, Springer, 2019.
- [28] S. Das and D. Chakraborty, "Conceptualizing fuzzy line as a collection of fuzzy points," *Inf. Sci.* 598, pp. 86-100, 2022.

Reconfigurable Digital Twins for an Industrial Internet of Things Platform

Eliseu Pereira^{*†}, Maria Arieiro^{*}, Gil Gonçalves^{*†}

[†]SYSTEC - Research Center for Systems and Technologies

^{*}Faculty of Engineering, University of Porto

Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

Email: {eliseu, up201606469, gil}@fe.up.pt

Abstract—Digital Twin (DT) emerged as an Industry 4.0 concept that reflects the behavior of physical entities and processes in the digital environment providing real-time information and insights. Typically, DTs are deployed on specific scenarios or components, virtualizing services and monitoring process variables. This extremely focused software implementation makes harder the adaptation and replication of DTs for new components with similar characteristics. This work's main goal was to develop a flexible platform that facilitates the implementation and modification of DTs for IEC-61499 compliant Cyber-Physical Systems (CPS), allowing their functionalities management and variables monitoring. The DT component was integrated into a web-based platform (Jurassic Park) that manages IEC-61499 compatible devices executing the runtime environment (DINASORE). The platform was validated in 2 different scenarios, 1) in a distributed system composed of different DTs and 2) controlling a gripper of a robotic arm.

Index Terms—Cyber-Physical Systems, Digital Twin, IEC-61499, Industrial Internet of Things, Virtualization.

I. INTRODUCTION

The changeover to Industry 4.0 was introduced in the shop-floor of new information and communication technologies such as the Industrial Internet of Things (IIoT), Cyber-Physical Systems (CPS), among others. The digitalization and virtualization of the physical devices provided to users and operators support for communicating with other devices, tracking errors, optimizing production and many other advantages [1][2]. Digital Twins (DTs) implement virtual models of physical entities to mirror the geometry and behaviour of that entities in the digital domain. With those digital models, real-time monitoring and control of the physical entities can be achieved [3]. Cyber-Physical Production Systems (CPPS) have a need to be easily reconfigurable, since these systems stand out by responding quickly to changes in production and being flexible enough to introduce new functionalities according to the different needs of the production lines. One of the existent standards for the CPPS reconfiguration is the International Electro-technical Commission (IEC) 61499 standard [4], which allows the encapsulation of different functionalities in software modules, the so-called Function Blocks (FBs).

Several DTs proposals have been made recently, however there is a lack of flexible solutions that are applicable in the production lines. Nowadays, the creation of a DT is a centralised process, oriented towards the device that is intended to be reflected in the digital world. This makes the

DT not very flexible. In the complex case of a system with a diversity of physical entities fulfilling different functionalities, reusing a DT is a critical process and will require increased computational and human effort. This limitation could be solved through the usage of standards for the deployment of reconfigurable CPPS, such as the IEC-61499 standard. However, its integration with the DT concept is still limited and barely explored by the scientific community. Some of the most relevant limitations are the difficulty to manage complex systems with many software modules and the lower flexibility and adaptability of the systems to new tools, at the software and hardware level [5].

The main objective of this work was to develop a flexible DT application, integrated with a web-application (Jurassic Park [6]) and with the physical devices (executing the DINASORE [7] runtime environment). The DT solution will include 1) a monitoring module, 2) a control module and 3) a visualisation module. The monitoring module is intended to be the processing component of the information that arrives from the physical environment. The main purpose of this module is to understand the behaviour of the physical entities and store that information in the digital world. The DT control module aims to be the response feature of the DT solution. Through this module, the user can act on the physical entities, thus making the system not only an analysis component but also an acting component with the purpose of modifying the system. The general idea behind the DT visualization module is to allow the management of the DTs from the application system. The solution was validated in 2 different scenarios, 1) used to monitor and control a distributed CPS, and 2) to control a robotic arm gripper.

The remainder of the paper is structured as follows: in Section II, we present the literature review, Section III exposes the implementation carried out and Section IV presents the main tests and results made to validate the DT solution developed. The conclusions and future work are presented in Section V.

II. LITERATURE REVIEW

Digital Twins (DTs) can adopt different methodologies, concepts and technologies, providing capabilities of simulation, monitoring or optimization. For that, we initially present the main technologies used as basis for the implementation of the

DT platform, and after the related work presents a description of the more relevant DT implementations.

A. Background

There are several enabling technologies that allow the implementation of CPSs [8], for different purposes, like device management, data collection, workflow orchestration, among others. The DT solution stacks different technologies, having as basis an environment with essential capabilities, like reconfiguration of software, data interoperability or communication transparency [9].

The DINASORE [7] is a distributed platform which enables the pre-processing of data in edge devices, using Python based FBs. This platform allows the construction of a wide variety of FBs by the user, with different goals, like to control drivers, integrate sensors, apply processing techniques, among other functionalities. The DINASORE uses the 4DIAC-IDE as graphical user interface to orchestrate the FB pipelines, deployed after in the distributed CPS.

The 4DIAC-IDE [10] is a tool developed by Eclipse that allows the orchestration of a workflow of FBs, after being deployed in a distributed system. With this tool, it is possible to draw the FBs distributed systems through a graphical interface and then deploy it into the devices connected to the network.

The Jurassic Park [6] is a web-based application that acts as a centralised repository of FBs, allowing the creation and management of FBs. The Jurassic Park is composed of two individual components, the backend and the web application. This application is integrated with the DINASORE and the 4DIAC-IDE, allowing the full deployment and management of distributed systems through the IEC-61499 standard.

B. Related Work

As identified in [11], currently there is a huge need in providing on-demand manufacturing services through Industrial Internet of Things (IIoT) networks. These services are easily achieved with the implementation of cloud-based manufacturing solutions. The literature review presented in [12] shows a set of cloud solutions available to help in the implementation of Digital Twins (DTs), in particular the Microsoft Azure IoT solution and the Amazon Web Services (AWS) IoT solution. When large industrial companies began to realize that the implementation and consequent application of DTs could have a huge impact on the efficiency of their production lines, they begin their own initiative to develop their own tools [13][14].

In [15], the authors present DT use cases in the industrial environment, where Siemens presents two DT implementations, one for the power system and the other for the wastewater plant. General Electric has not only developed a DT for a wind farm, but has also been able to prove through this implementation that a wind farm should be operated, developed and maintained in a more efficient way. British Petroleum has developed a DT of oil and gas facilities located in more restricted areas. The air vehicle manufacturer Airbus has used DT-based solutions to monitor its production lines and to optimise its operations. Going into a more specific example

of a DT implementation in the industry context, [16] presents a DT within a production line of hollow glass manufacturing. The authors of this implementation argue that this work has allowed them to realize that digital simulations have a great impact on how production lines can be optimized, since the simulations reflect their behavior in the real world.

In [17], the authors present a DT solution in the commercial greenhouse sector. The article highlights the need of the commercial greenhouse production systems to become more energy-efficient while maintaining a sustainable production. Therefore, the DT solution mentioned was proposed in order to provide a control and monitoring feature of the production flow of a greenhouse production system. The authors of this article also defend two essential ideas. One of them is that, with DT implementations, it is possible to turn the industrial environment more energy and climate friendly and that leads to a reduction of costs within the production lines. The other idea is that, nowadays, the industrial environment requires developed generic DT frameworks capable of controlling the processes accurately and to respond to changes in the orders of the production lines.

III. IMPLEMENTATION

The DT platform was conceptualized to structure multiple DTs, allowing their monitorization and control. This concept is composed of three main objects, 1) the digital twin, that is at the top of the architecture and is responsible for controlling the functionalities of the physical equipment, 2) the functionality, which describes the DT characteristics, mainly the monitoring system variables and the functions to trigger the execution of certain actions, and 3) the device, that reflects the physical entity in the digital world, like sensors, machines, etc. The DT platform developed is embedded in the Jurassic Park taking advantage of some of its features. Figure 1 shows the DT proposed solution composed of three components, the DT visualisation module, the DT monitoring module and the DT control module.

The DT visualisation component (presented in Section III-A) allows the interaction between the DT platform and the user. This component is incorporated into the web application of the Jurassic Park. The main functionalities of this visualisation component focus on the management of the DTs. Additionally, this component will be responsible for triggering the monitoring and control requests from the DTs. The DT monitoring module (presented in Section III-B) is intended to serve as a processing tool of the information that arrives from the physical entity to the digital one and serves the information to the user through the graphical interface. The component is incorporated into the web server of the Jurassic Park using the Open Platform Communications - Unified Architecture (OPC-UA) communication protocol as data source to collect the information from the physical world. The monitoring module collects data from the FBs variables and events that are running on the DINASOREs connected to the platform, allowing the user to filter the ones to monitor. The DT control component (presented in Section III-C) permits to trigger functionalities

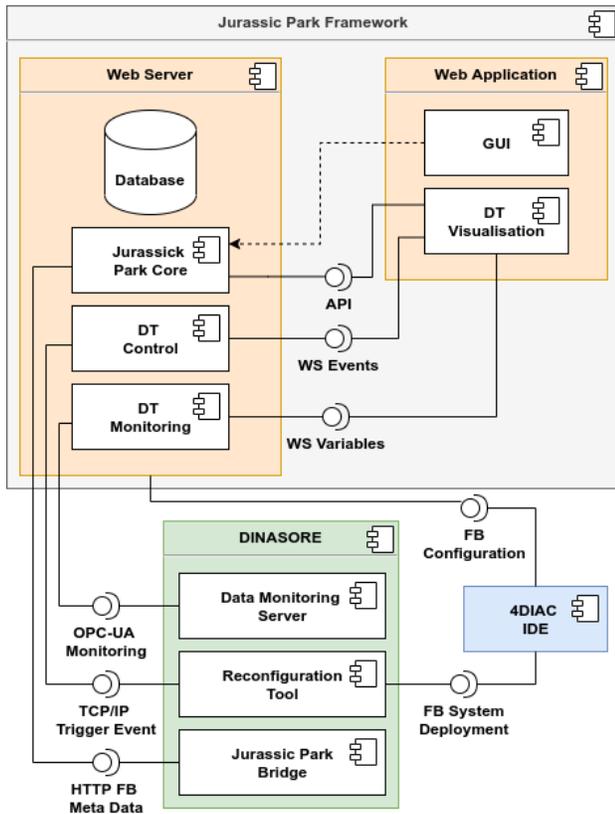


Fig. 1: DT architecture integrated with the Jurassic Park and DINASORE.

over the physical entities connected to the platform. The component uses the variables and events from the FBs that are running on the devices connected to the platform and assigns them a trigger feature. In this way, the user will have the opportunity to manipulate the events of the FBs. As in the case of the monitoring component, the control component is also incorporated into the web server.

The backend component (web server) is a complex module that uses several technologies to consolidate the integration of all the services it supports. This component is responsible for the HTTP API that manages the FBs and the real-time communication of the devices connected to the platform as well as the database. One of the most relevant technologies used in the backend is the Node.js framework, which integrates several components of the backend such as the API, the real-time communication with the web application and the real-time communication with the DINASORES, through the OPC-UA communication protocol. The backend component uses as database the MySQL database. The backend includes also an File Transfer Protocol (FTP) server that makes available the download of FBs by the DINASORES.

The frontend component (web application) is the graphical user interface of the platform, and it was developed using React [18]. This component is responsible for the page components (buttons, boxes, tables, among other User Interface (UI) elements), which are customized according to the existent

resources in the CPS. The automatic update of the information is performed using the Socket IO library, which supports real-time communication with the backend, and avoids the need of refreshing the web page to update the information.

A. Digital Twin GUI

The DT visualisation component allows the user to create and manipulate DTs. The platform allows the user to group the devices they want to monitor and control into a category called DT and subsequently they can associate that DT with a general functionality. Therefore, in order to describe the page elements in a more user-focused way, we list below the different pages implemented to fulfil the requirements of the DT component.

The new DT page allows the user to add new DTs with a specific name and to associate them with respective DINASORES. The page has a menu for creating a new DT, and in this menu the user has access to all the devices that are communicating with the Jurassic Park in real-time. Therefore, to create a new DT, the user only has to complete the field to insert the name of the new DT and open the list of available devices and selected those he wants to associate. It should be noted that the user can choose more than one DINASORE. Additionally, the user can only create the new DT if he has chosen at least one DINASORE.

The Digital Twin monitoring page, in Figure 2a, performs all the management of the DT. On this page, the user can observe the functionalities that are currently active with the respective DT. Besides that, it is also on this page that the user can create other functionalities and associate a DT capable of observing the variables and/or events of interest. Looking first at the feature of the page concerning the creation of a new functionality, the user can choose the name of the functionality they want to add to the monitoring platform. After choosing the name, the user will also have to choose one DT from the range of available DTs that were previously created in the New Digital Twin page, which they want to associate with the new functionality. Through this page, the user will have at his disposal a table, listing all the functionalities currently active and a set of features. The *Details* feature allows the user to view in detail the information that is being collected from the variables and events that are currently being monitored. The *Add Details* feature enables the user to choose the variables and events that they want to monitor on a given FB. With the *Edit* feature, the user can edit the name of the functionality. In addition to being able to edit and manipulate the variables and events that each functionality will monitor, the user can also delete the available functionalities they want with the *Delete* feature.

Whenever the user presses the functionality details button, the interface redirects to a new page. The functionality details page, in Figure 2b, has two different tables, one for monitoring variables and another for triggering events. Note that the page only displays the variables and events that have been previously selected on the DT monitoring page. Thus, the user has at their disposal information organised either by variables or by events. Regarding the variables monitoring,

| Functionality ▲ | Digital Twin | Details | Add Details | Edit | Delete |
|--------------------|-----------------------|---------|-------------|------|--------|
| Gripper | DT_Gripper_test | ⚙️ | ⊕ | ✎ | 🗑️ |
| Optimization_test | DT_test_optimization | ⚙️ | ⊕ | ✎ | 🗑️ |
| Sensorization_test | DT_test_sensorization | ⚙️ | ⊕ | ✎ | 🗑️ |

New Functionality

Insert new functionality name *

ADD FUNCTIONALITY

(a) DT monitoring web page.

Optimization_test

| Variable | Function Block | Smart Component | Current Value | Delete |
|-------------|-----------------|-----------------|--------------------|--------|
| COST | ENERGY_COSTS_1 | dinasore3 | 1269.8766074325235 | 🗑️ |
| TEMPERATURE | OPTIMIZE_ENERGY | dinasore3 | 1 | 🗑️ |

| Event | Function Block | Smart Component | Trigger Event | Delete |
|-------|----------------|-----------------|---------------|--------|
| READ | ENERGY_COSTS | dinasore3 | ⚡ | 🗑️ |

(b) Functionality details web page.

Fig. 2: Digital Twin graphical interface, including the monitoring (a) and functionalities details (b) web pages.

the user can observe in the table some information of the variable such as the variable name, the FB where the variable is being monitored, the device that is allocating and also the current value of the variable. Additionally, the user can delete the variables that they no longer wish to monitor by clicking on the delete button present in each row of the table. The event monitoring table does not differ much from the variable monitoring table. It is also possible to see the name of each event selected by the user, the associated FB and the DINASORE where it is being mapped. However, unlike the variable monitoring table, the event table has a button that allows the user to trigger an action on the associated event. When the user presses the trigger event button, the Jurassic Park automatically sends a request to the server to execute the event on the associated FB. The user also has a button available on the event table that allows them to delete events that they no longer wish to observe.

B. Digital Twin Monitoring

The monitoring component of the DT platform is responsible for requesting and collecting information on each DT, which is then made available in the DT visualisation component. For that, the user should create the different DTs and functionalities, as well as define the monitored devices, variables and events. With the previous objects created, the next integration is to automatically receive real-time feedback of the values of the monitored variables. Initially, the request is sent via WebSockets from the web application, where a persistent listener waits for feedback from the monitored

variables. On the server side, the communication with the DINASORES is performed using a OPC-UA client. When the monitored variables present a change, the OPC-UA client automatically identifies it, and then the data is sent back to the backend controller which ensures the notification of the new value to the web application, waiting for the information through the listener function. When the user no longer wants to observe the variables, an event to cancel the subscriptions is sent to the OPC-UA client.

C. Digital Twin Control

The DT control component allows the user to trigger events in certain FBs running on specific DINASORES. After interacting with the UI element, a particular message, containing the information relating to the triggered event, is sent to the backend component. Having the information regarding the event in the backend, a function sends the action to the device in question, using Transmission Control Protocol/Internet Protocol (TCP/IP) sockets. For that, the function uses as input the event information, in particular the FB name, event name, device IP address and port. When the DINASORE receives the message, it pushes an event on the specified FB and executes the triggered functionality.

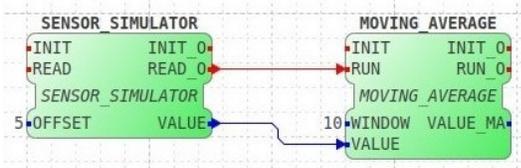
IV. EXPERIMENTS AND RESULTS

The experiments performed to validate the DT component of the platform focused on two use cases that allow to test the different implemented features. The experimented use cases are 1) the monitorization of simulated sensors and

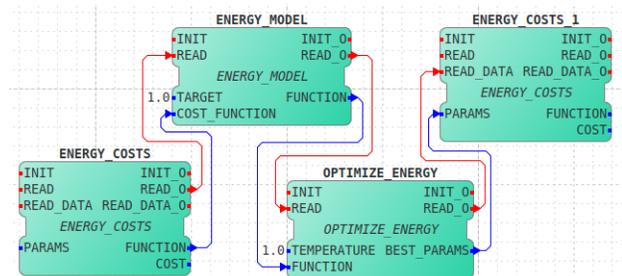
the optimization of energy in a distributed CPS, and 2) the manipulation of a robotic arm gripper.

A. Monitoring and optimizing a distributed CPS

The first validation scenario is a distributed system with different devices connected to the platform. This experiment consisted of connecting a set of raspberry pis to the platform, which are responsible to execute two different FBs pipelines/workflows, one for sensing purposes and the other for energy optimisation. This set of FBs pipelines allows the monitoring of variables and trigger of events in a distributed system, allowing the validation of the previously explained monitoring and controlling features of the DT platform.



(a) Sensor simulation FBs workflow.



(b) Energy optimization FBs workflow.

Fig. 3: Distributed CPS composition with a sensorization (a) and optimization (b) workflows.

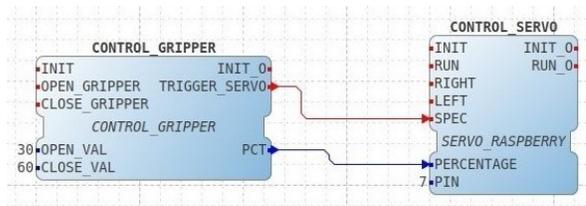
The first raspberry pi was used to validate the variable monitoring component, using a small FBs workflow, that simulates a sensing system, present in Figure 3a. The FBs composing the workflow are 1) the SENSOR_SIMULATOR, responsible for generating a random value to simulate the value of a sensor, and 2) the MOVING_AVERAGE, which calculates the average of the last N values. After the deployment of the workflow, the next step was to create a DT for the sensing component, associated with the raspberry pi that executes the DINASORE. Then, we created a functionality that includes the following monitoring variables, 1) the variable VALUE (variable containing the simulated sensor data) from the SENSOR_SIMULATOR FB and, 2) the variable VALUE_MA (variable containing the moving average which is calculated) from the MOVING_AVERAGE FB.

The second raspberry pi hosts a set of FBs that optimize the energy consumption of a process, present in Figure 3b. These FB pipeline allows the validation of the event triggering process through the DT platform and consequently the monitoring of variables linked to that event. The FBs composing

the workflow are 1) the ENERGY_COSTS, which specifies the function for energy costs regarding velocity and power, 2) the ENERGY_MODEL, which is the model that allocates the energy consumption, using as input the energy costs, and 3) the OPTIMIZE_ENERGY, which optimizes the given function, through the use of the Dual Annealing algorithm. The DT associated to the optimization workflow will include a functionality that contains the following variables and events: 1) the COST variable (optimised energy costs) present in the ENERGY_COSTS_1 FB, and 2) the READ event (triggering optimisation) of the ENERGY_COSTS FB. Thus, with this variable and event combination, we are able to trigger an event via the DT platform, sent to the raspberry pi, that executes the optimization.

B. Creating a Digital Twin for a Gripper

The second validation scenario focuses on the creation of a DT for a robotic arm gripper, validating the control and monitoring features in a physical component. The component used was a 3D printed gripper using a raspberry pi as controlling system for the servo motor. This experiment consisted in controlling the gripper through the developed DT platform by manipulating the event trigger to make the gripper arm open and close according to the event, as shown in Figure 4b and 4c. The FB workflow that allows the control of the gripper, present in Figure 4a, is composed of the FBs: 1) the CONTROL_GRIPPER, responsible for identifying whether there is a request to open or close the gripper, depending on which, the FB sends the corresponding percentage to the output (PCT), and 2) the CONTROL_SERVO, which receives as input the percentage and then updates the general-purpose input/output (GPIO) that controls the servo motor with the corresponding value.



(a) FBs workflow to control the gripper.



(b) Gripper open.

(c) Gripper closed.

Fig. 4: Results obtained with the gripper scenario, including the FBs workflow (a) and the physical gripper (b) and (c).

The DT created intends to monitor and control the gripper manipulation. For this use case, two distinct events and a variable (to evaluate whether the events are being properly triggered) have been added to the functionality associated to the DT. The selected events to control the gripper were the OPEN_GRIPPER and CLOSE_GRIPPER events. In order to check whether the control on the DT platform is functional, the gripper would have to open when the user pressed the release button. If the close button was pressed, the gripper would have to execute the closing movement. The execution of both movements were validated visually and also through the monitoring variable (percentage), which confirms if the gripper moves according with the monitored percentage.

V. CONCLUSION

In conclusion, the main objectives of this platform were to implement a flexible and reconfigurable DT solution capable of increasing the monitoring capacity and enable the remote control of a CPS. The components developed, mainly the DT visualisation component, the DT monitoring component and the DT control component permitted the CPS to accomplish the described attributes, like DT reconfiguration. The DT solution was integrated together with a mature stack of technologies including the Jurassic Park and the DINASORE. The experiments performed support the usability and flexibility of the web based platform. Finally, it is important to highlight that one of the greatest contributions of this project was to develop a platform that, given its flexibility and configurability, can be easily integrated in the industrial sector and supports the IEC-61499 standard.

As future work, one of the main goals will be the storage of data generated by the DT variables, taking into consideration constraints of data volume, in terms of storage and data flow/rate. On top of this unstructured database, it will be possible to implement predictive algorithms to forecast and optimize the behavior of DTs. Additionally, the entire platform (Jurassic Park with DT component and DINASOREs) will be integrated in an industrial scenario composed of different machines, including integration of sensors, simulation of processes, and optimization of resources.

ACKNOWLEDGMENT

INDTECH 4.0 - New technologies for intelligent manufacturing. Support on behalf of IS for Technological Research and Development (SI a Investigacao e Desenvolvimento Tecnológico). POCI-01-0247-FEDER-026653

REFERENCES

- [1] G. Gonçalves, J. Reis, R. Pinto, M. Alves, and J. Correia, "A step forward on intelligent factories: A smart sensor-oriented approach," in *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, 2014, pp. 1–8.
- [2] L. Antão, R. Pinto, J. Reis, and G. Gonçalves, "Requirements for testing and validating the industrial internet of things," in *2018 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, 2018, pp. 110–115.

- [3] K. Ding, F. T. S. Chan, X. Zhang, G. Zhou, and F. Zhang, "Defining a Digital Twin-based Cyber-Physical Production System for autonomous manufacturing in smart shop floors," *International Journal of Production Research*, vol. 57, no. 20, pp. 6315–6334, 2019.
- [4] K. Thramboulidis, "Iec 61499 in factory automation," in *Advances in Computer, Information, and Systems Sciences, and Engineering*, K. Elleithy, T. Sobh, A. Mahmood, M. Iskander, and M. Karim, Eds. Dordrecht: Springer Netherlands, 2006, pp. 115–124.
- [5] G. Lyu and R. W. Brennan, "Towards IEC 61499-Based Distributed Intelligent Automation: A Literature Review," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2295–2306, 2021.
- [6] J. Pedro Furriel, E. Pereira, J. Reis, and G. Gonçalves, "Jurassic park - a centralized software modules repository for iot devices," in *2021 10th Mediterranean Conference on Embedded Computing (MECO)*, 2021, pp. 1–4.
- [7] E. Pereira, J. Reis, and G. Gonçalves, "Dinasore: A dynamic intelligent reconfiguration tool for cyber-physical production systems," in *Eclipse Conference on Security, Artificial Intelligence, and Modeling for the Next Generation Internet of Things (Eclipse SAM IoT)*, 2020, pp. 63–71.
- [8] Q. Qi, F. Tao, T. Hu, N. Anwer, A. Liu, Y. Wei, L. Wang, and A. Nee, "Enabling technologies and tools for digital twin," *Journal of Manufacturing Systems*, vol. 58, pp. 3–21, 2021, digital Twin towards Smart Manufacturing and Industry 4.0.
- [9] E. Pereira, R. Pinto, J. Reis, and G. Gonçalves, "Mqtt-rd: A mqtt based resource discovery for machine to machine communication," in *Proceedings of the 4th International Conference on Internet of Things, Big Data and Security - IoTBDS*, INSTICC. SciTePress, 2019, pp. 115–124.
- [10] A. Zoitl, T. Strasser, and A. Valentini, "Open source initiatives as basis for the establishment of new technologies in industrial automation: 4diac a case study," in *2010 IEEE International Symposium on Industrial Electronics*, 2010, pp. 3817–3819.
- [11] Y. Lu and X. Xu, "Cloud-based manufacturing equipment and big data analytics to enable on-demand manufacturing services," *Robotics and Computer-Integrated Manufacturing*, vol. 57, pp. 92–102, 2019.
- [12] T. Catarci, D. Firmani, F. Leotta, F. Mandreoli, M. Mecella, and F. Sapio, "A conceptual architecture and model for smart manufacturing relying on service-based digital twins," in *Proceedings - 2019 IEEE International Conference on Web Services, ICWS 2019 - Part of the 2019 IEEE World Congress on Services*. IEEE, Piscataway, NJ, USA, 2019, pp. 229–236.
- [13] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open Research," *IEEE Access*, vol. 8, pp. 108 952–108 971, 2020.
- [14] A. Rasheed, O. San, and T. Kvamsdal, "Digital twin: Values, challenges and enablers from a modeling perspective," *IEEE Access*, vol. 8, pp. 21 980–22 012, 2020.
- [15] F. Tao, H. Zhang, A. Liu, and A. Y. Nee, "Digital Twin in Industry: State-of-the-Art," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019.
- [16] H. Zhang, Q. Liu, X. Chen, D. Zhang, and J. Leng, "A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line," *IEEE Access*, vol. 5, pp. 26 901–26 911, 2017.
- [17] D. Anthony Howard, Z. Ma, J. Mazanti Aaslyng, and B. Nørregaard Jørgensen, "Data architecture for digital twin of commercial greenhouse production," in *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, 2020, pp. 1–7.
- [18] A. Fedosejev, *React.js essentials*. Packt Publishing Ltd, 2015.

Applying Deep Learning Techniques in Automated Analysis of Echocardiograms, CMRs and Phonocardiograms for the Detection and Localization of Cardiac Diseases

Arvind K. Bansal and Racheal Mukisa

Department of Computer Science
Kent State University
Kent, OH 44242, USA

email: akbansal@kent.edu and rmukisa1@kent.edu

Abstract— Echocardiography (echoCG), Cardiac Magnetic Resonance (CMR) and phonocardiograms (PCG) are becoming indispensable tools in the diagnostics and management of cardiac diseases due to advancements in imaging techniques, improvement in processing power, availability of large multimedia databases in Electronic Medical Records (EMR) and rapid lowering of cost. Image-based and video-based data in echoCG and CMR are multi-dimensional and exceed the capabilities of traditional statistical learning. Deep learning technologies provide new possibilities for accurate, consistent, and automated interpretation of echoCG, CMR, and PCG, reducing the risk of human error. Deep learning and signal analysis techniques are being applied to analyze these complex data for improved diagnosis of cardiac diseases involving heart muscles, valvular defects, cardiac chamber deformities negatively affecting blood-oxygenation and blood-flow. This review describes applications of deep learning techniques, such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-term Memory Neural Network (LSTM), transfer learning, and their variations in enhancing the classification of heart diseases from echoCG, CMR and PCG.

Keywords—Artificial Intelligence; CMR; deep learning; echocardiogram; heart diseases; machine learning; Phonocardiogram.

I. INTRODUCTION

The importance of Cardiovascular Diseases (CVD), Congenital Heart Diseases (CHD) and other heart-related diseases is significant. According to World Health Organization (WHO), 17.9 million persons died annually of CVD worldwide [1]. Total prevalence of cardiac diseases in 2018 was around 126.9 million USA alone [2]. Infants are born with CHD with a mean rate of 8.2 per thousand births per year worldwide [3]. Their long-term prognosis for survival is low. Invasive surgery is expensive, risky and not advisable for diagnostics and maintenance of heart-conditions post-surgery.

Electrocardiogram (ECG), echocardiogram (echoCG), Cardiac Magnetic Resonance (CMR) and digital phonocardiograms (PCG) are minimally invasive techniques that have been used to assess heart abnormalities [4]-[11]. Computer Tomography (CT) is also used to study heart

defects [8]. However, CT scans are associated with radiation and have lesser resolution than CMR [8].

ECG is the least invasive technique and suitable for indicating CVD affecting change in emitted waveforms. However, it is not well suited to assess localization of structural defects and motion-related deformities in heart-muscles (such as hypertrophic cardiomyopathy) and heart-valves (such as mitral valve regurgitation or aortal stenosis) and cannot assess blood-volume flow – an important feature to assess the heart diseases due to smaller volume of blood-flow, turbulent blood-flow, stenosis, mixing of oxygenated and deoxygenated blood due to holes in the septum, and ischemia caused by plaque formation and arteriosclerosis.

Due to improved resolution in inexpensive echoCGs, CT scans are less preferred and have been left out in this review. ECG analysis has been left out due to its limitations in localization of cardiac defects. In this paper, we review the applications of deep learning methods to analyze echoCG, CMR and PCG.

Recent progress in echoCG has made it quite accurate, inexpensive, and a preferred alternative for assessing structural and blood-flow-related diseases. CMR has the highest resolution. However, it is expensive. EchoCG is preferred to measure the speed of blood-flow and blood-turbulence present in many valvular diseases [4]. PCG is analyzed to diagnose limited cardiac diseases based upon emitted sound, while echoCG and CMR use images, image-sequences and video-clips of heart-muscles and valves [7][9]-[13].

Guidelines have been developed to ensure accurate interpretation of echoCG, CMR and PCG outputs [12]. However, the final analysis heavily relies on the operators' experience and knowledge. This causes subjectivity and variability in interpreting image and sound patterns [12].

In the last decade, image analysis and speech-recognition has improved significantly due to rapid advances in deep learning techniques, such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-term Neural Network (LSTM), transfer learning, and their variations and combinations. Deep Neural Networks (DNN) exploit convolution-filter-based feature-maps, temporal context, memory, and selective forgetfulness in the artificial neurons to improve object-classification [14]-[17].

AI and deep learning techniques are being exploited to generate accurate, consistent and automated interpretation of echoCGs, CMR and PCG to diagnose cardiac diseases related to structural and heart muscles related defects, valvular defects, and blood-flow-related diseases [4]–[6]. For example, machine learning models have been shown to provide an instantaneous assessment of the left ventricular ejection fraction and longitudinal strain [13].

Large volumes of multidimensional imaging data generated in 2D, 3D and 4D (includes 3D images and temporality present in video clips) formats are available (see Section VII). EchoCG requires image-classification and video-analysis; CMR requires analysis of 2D and 3D cardiac images and 4D videos. PCG requires waveform segmentation and analysis to classify the wave-patterns.

Many factors have contributed to the recent interest and the advancement of cardiac image and motion analysis: 1) multifold improvement in deep learning techniques in the last decade; 2) availability of large scale of high-resolution CMR and echoCG datasets; 3) lowering of the cost of CMR and echoCG; 4) recent acceptance of automated intelligent image analysis techniques by the clinicians in other disease domains such as mammograms; and 5) cost and side-effects of invasive exploration for cardiac diseases.

The rest of this paper is organized as follows. Section II describes the background concepts: heart anatomy, echoCG, CMR, and deep learning techniques. Section III describes diseases associated with defects and deformities in heart muscles, valvular (heart-valves related) and blood-flow-related diseases. Section IV describes echoCG analysis. Section V describes CMR analysis. Section VI describes PCG analysis. Section VII describes major data repositories. The last section concludes the discussion.

II. BACKGROUND

A. Heart Anatomy

A heart (see Figure 1) comprises four chambers: Left Atrium (LA); Right Atrium (RA); Left Ventricle (LV); Right Ventricle (RV). LA gets oxygenated blood from the lungs; LV pumps oxygenated blood to the body; RA collects deoxygenated blood from the body; RV sends deoxygenated blood to the lungs. The left and right sides are separated by a muscle called ‘septum’ [5]. A hole in the septum contaminates oxygenated blood in the LV (or LA) with deoxygenated blood in the RV (or RA).

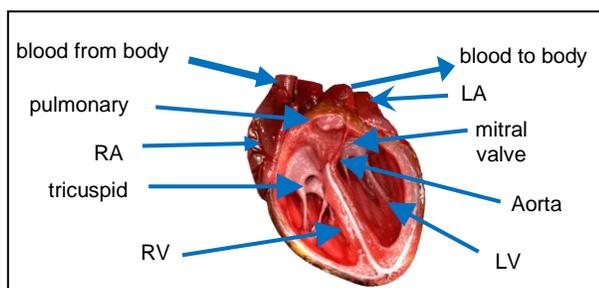


Figure 1. An illustration of heart anatomy [18].

A heart has four valves: 1) *aortic valve* regulating blood-flow from LV to the body; 2) *mitral valve* regulating blood-flow from LA to LV; 3) *pulmonary valve* regulating blood-flow from RV to lungs; 4) *tricuspid valve* regulating blood-flow from RA to RV. Valves comprise two or three leaflets that open and close synchronously. Periodic relaxation and compression of cardiac chambers cause blood flow. Two valves, *aorta* and *mitral*, play key roles in the oxygenated blood-flow from the heart to the body.

B. Echocardiogram

EchoCG (see Figure 2 [18]) is an ultrasound-based technique that assesses the reflectivity and refraction of emitted microwaves altered by tissue type and density [5]. There are two types of echograms: *transthoracic echoCG* (TTE) and *transesophageal echoCG* (TEE). TTE is a noninvasive and preferred technique [19].

By comparing the differences between signal reflected between the healthy tissue and the query tissue, the technique assesses the presence and the extent of unhealthy tissue layers and their thickness [8]. By combining echoCG with Doppler effect, blood-flow speed and direction are also estimated [16].

C. Cardiac Magnetic Resonance (CMR)

Cardiac Magnetic Resonance (CMR) is a high-resolution imaging technique using a strong magnetic field that excites hydrogen ions in water molecules inside tissues in a Region of Interest (ROI). CMR measures the emitted energy when hydrogen ions return to the normal state [8]. The relaxation time differs between healthy tissues and diseased tissues. By knowing the relaxation patterns of tissues, CMR images are formed. Pixels/voxels can be imaged [8] in an ROI by superimposing non-uniform magnetic-field with a static strong magnetic field,

Computational combinations of multiple slices of CMR in images give a 3D CMR image that is visualized using computer graphics. As shown in Figure 3 [21], CMR is used to analyze the deformities in different cardiac chambers, and changes in their motions and blood-flow patterns [8]-[11].

D. Phonocardiogram (PCG)

Phonocardiograms are high-fidelity sound recordings (see Figure 4 [22]) generated by continuous opening and closing of cardiac valves, and blood-flow through arteries in heart-chambers [7][23][24]. Different valvular diseases create different sound-patterns (see Figure 4) helping in the classification and identification of valvular diseases.

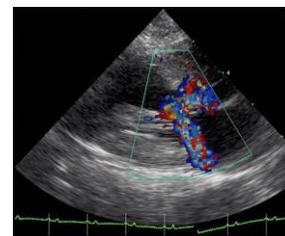


Figure 2. An illustration of echocardiogram [20].

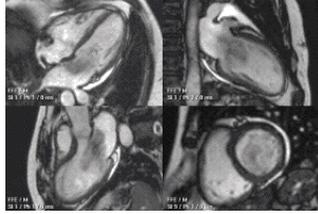


Figure 3. An illustration of CMR image [21].

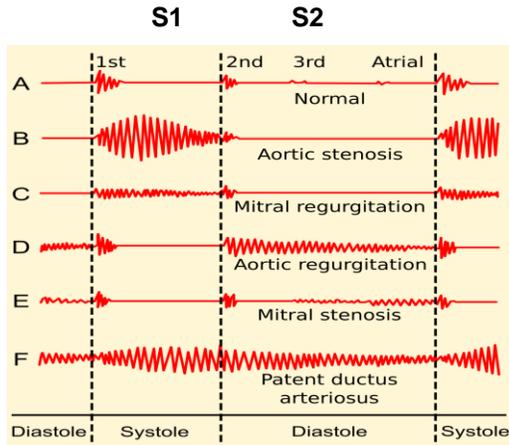


Figure 4. Phonocardiogram of different heart defects [22].

PCG comprises two major sound patterns S1 and S2, for each cardiac cycle, as shown in Figure 4. The sound S1 occurs at the beginning of the ventricular systole due to the closure of atrioventricular valves: *mitral* and *tricuspid*. The sound S2 marks the beginning of ventricular diastole and comprises two components due to the closure of *aortic* and *pulmonary* valves [7][23].

III. VALVULAR AND HEART MUSCLE DISEASES

Classes of cardiac disease derived using deep learning techniques are: 1) deformity in the heart-structure or the presence of a tumor; 2) calcification causing a lack of synchronization in heart-valves resulting in blood-leak (stenosis), or regurgitation in the corresponding blood channels; 3) thickening of cardiac walls in one or more chambers reducing the blood-flow volume and negatively affecting the heart-motion; 4) plaque formation in an artery or vein restricting blood-flow [4][9][13][19].

Deformities in the valves cause *stenosis* - restricting the blood-flow, or *regurgitation* - turbulent blood-flow and/or blood-flow in reverse direction around a valve. Four classes of valvular diseases have been computationally analyzed using DNN in recent years [4][19]. The diseases are *aortic stenosis*, *mitral stenosis*, *aortic regurgitation*, and *mitral regurgitation*. Uneven calcifications on valve-leaflets cause a lack of synchronization in the opening and closing of valves resulting in regurgitation and stenosis [4].

Another valvular disease is *Rheumatic Heart Disease* (RHD) that occurs as an after-effect of rheumatic fever, resulting in valvular lesions in aorta valve, mitral valve or tricuspid valve, weakening the valve-function. Affected

valves develop regurgitation and stenosis, with the most common being *mitral regurgitation* [4][19].

Another class of disease is *Congenital Heart Diseases* (CHD) where the cardiac structure has defect(s) since birth [18][19]. Common CHDs are: 1) *Congenital Valve Disease* (CVD), including *Bicuspid Aortic Valve Disease* (BAVD), *Pulmonary Valve Stenosis* (PVS), *Aortic Valve Stenosis* (AVS), and *Ebstein’s anomaly*; 2) *Atrial Septal Defect* (ASD) – a hole in the septum between upper chambers connecting LA and RA; 3) *Ventricular Septal Defect* (VSD) – a hole in the septum in lower chambers connecting LV and RV; 4) *Coarctation of the aorta* (CA) – narrowing of the aorta after it leaves the heart causing blood-turbulence; 5) *Patent Ductus Arteriosus* (PDA) - connection between two blood channels because channels do not close after birth (see Figure 5 [26]).

Ebstein’s anomaly is a severe heart defect in leaflets of a *tricuspid valve* (between RA → RV) restricting blood-flow to lungs, resulting into the lack of oxygenation. *Tetralogy of Fallot* is a combination of four CHDs: *VSD*, *PVS*, a *misplaced aorta* and *right ventricular hypertrophy*. It causes the lack of blood-flow and deoxygenation of blood, resulting in ischemia and vascular degeneration.

Other muscle-related conditions are *cardiac myopathy* (thickening of heart muscles) and *cardiac hypertrophy* (smaller volume in heart chambers due to the thickening of cardiac walls) [8]. *Myopathy* causes heart chambers to contract slowly, reducing the blood-flow [13][19]. *Cardiac hypertrophy* is accompanied by *myopathy* [8]. *Cardiac hypertrophy* is identified by analyzing the wall-thickness combined with the ejection-volume estimation from the LV.

Applications of deep learning techniques for diagnosing cardiac muscles-related and valvular diseases are summarized in Table I.

IV. ECHOCARDIOGRAM ANALYSIS

EchoCG analysis can diagnose valvular diseases, such as *stenosis* and *regurgitation*, *atrial blockages*, *atrial fibrillation*, *congenital heart disease*, *coronary arterial disease*, *cardiomyopathy*, *cardiac hypertrophy*, and *murmur* [19][25][28]-[32].

TABLE I. DNN LEARNING BASED ANALYSIS FOR CARDIAC DISEASES

| Disease Class | Input Mode | DNN Technique |
|-------------------------------------|----------------------|--|
| Valvular stenosis + regurgitation | echoCG and PCG | CNN-based segmentation and TGNN + CNN + LSTM |
| Fetal heart defects | echoCG + CMR | CNN-based segmentation for wall boundaries |
| Myocardium hypertrophy and myopathy | Doppler echoCG + CMR | Hybrid CNN + LSTM + encoders and decoders for wall thickness, chamber boundaries and blood flow volume |
| Ischemia and myocardial infarction | CMR | CNN based tissue classification |

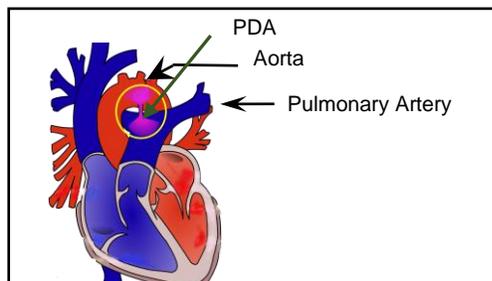


Figure 5. An illustration of PDA [26].

Aortal and mitral valve regurgitations have been identified using R-CNN (region-based CNN) on color Doppler echoCG [19]. R-CNN uses a bounding rectangular box around the objects and uses a combination of CNN and Support Vector Machine (SVM) for classification and object-detection [17]. Regurgitation is estimated semi-quantitatively using jet-area ratios. The area of an orifice is measured using ‘proximal flow convergence’ of blood-flow measured using Doppler effect [32].

Segmentation for valvular regurgitation uses multiple techniques, such as non-linear anisotropic diffusion [8][32]. *Regurgitation* is estimated using *Proximal Isovelocity Surface Area (PISA)*, which occurs when fluid flows through a circular orifice [33]. As fluid passes through a narrow orifice, it speeds up that can be measured.

CHDs have been identified using multiview (five views) echoCG analysis using multichannel CNN [25]. The advantage of having multiple views is clarity and supporting evidence from other views [19][25]. To reduce overfitting due to the limited availability of data, *depth-wise separable convolution* has been used. A standard convolution filter combines inputs in the same step. *Depth wise convolution* filters and combines in two separate layers, reducing the computational complexity.

Video of each view is fed to an encoder for frame-independent feature representation. It uses RNN to assign different weights to frames with the highest weight to the most relevant frame and progressively reduces weights for its neighbors. It uses *temporal convolution* to identify the neighboring frames [34]. *Temporal convolutions* are used on time-series data to maintain temporality and causality.

Common fetal heart defects are: (1) *tetralogy of Fallot*; (2) *ventricular dysplasia*; (3) *Ebstein’s anomaly*. The major problem in a fetal heart defect is the malformation of a subset of cardiac chambers and the leakage between the chambers due to holes in the septum. Major problems in the identification of cardiac chambers from infants’ echoCGs are: 1) artifact; 2) discrimination between chambers; 3) missing boundaries. The physical boundaries between atrium and ventricles are missing when the corresponding valves are opening. This requires deep learning techniques for echoCG analysis to identify four chambers and their motions accurately using segmentation.

Researchers have developed a cascaded dual layered Deep Wide Neural Network (DW-net) for semantic

segmentation of echoCGs [35]. DW-net comprises two layers: (1) Dilated Convolutional Chain (DCC) and W-net. DCC collects local and global features for the localization of ROI. W-net derives precise boundaries in ROI by repeated applications of encoders and decoders.

EchoNet, a CNN based model, detects cardiac structure and anatomy, blood-flow volume, the presence of pacemakers, LV hypertrophy, enlargement of LA chamber along with the prediction of age, sex, and gender of a patient [33]. It is trained on a large feature-set comprising age, sex, weight, BMI, presence of a pacemaker, LA enlargement, LV hypertrophy, End Systolic Volume (ESV), End Diastolic Volume (EDV), and Ejection Fraction (EF) [33].

V. CMR ANALYSIS

CMR imaging is a reference for imaging due to higher resolution. The cardiac region is identified knowing that heart movements change voxel-intensities in the dynamic part of a heart. LV, being the largest moving chamber, has been targeted for segmentation. LV also circulates maximum blood volume. After a ROI is identified, segmentation techniques are applied to derive the LV. After identifying the LV region, model-based techniques, augmented with deep learning techniques, derive other chambers and boundaries.

CMR analysis is also applied to estimate EDV, ESV, EF, and Myocardial Mass (MM) [36][37]. This quantification has been used to study blood ejection from each chamber, especially LV, and has been used to estimate the extent of cardiac myopathy and cardiac hypertrophy [37][38].

A major problem in CMR analysis is the segmentation of various boundaries in the heart to detect chambers, valves, artery, septum, muscles and valvular defects, especially in the presence of variations of pathologies in different patients. Segmentation approaches are image-driven or model-driven [37]. The image-driven approach uses intensity-based histogram analysis and thresholding, clustering, region growing, pixel/voxel level classification and active contours to identify blood-pool, myocardium or appendages. The model-based approach uses statistical analysis to derive atlas or shape contours.

The segmentation methods are classified as pixel-level classification, graph-based methods, probabilistic models, such as Markov Random Fields, deformity-based models, atlas-based methods, CNN and LSTM based deep learning models, and hybrid techniques that integrate deep learning techniques and deformity-based models [39][40]. Deformity-based models are based upon shape estimation. However, pixel-level classification, shape-based models, including deformity-based models (shape contours), and probabilistic models suffer from inaccuracies [40]. Random forest models are based on intensity classification that causes inaccuracy. Combined with deep learning models, deformity-based models accurately estimate the dynamic shape of various chambers [40].

LV segmentation has been used to derive motion estimation, wall thickness on different sides, local deformation and myocardial strain during systole and diastole phases [37]. Wall thickness changes for an ischemic heart [38]. Wall thickness in the LV during the systole phase

is used to estimate malfunction in myocardial ischemia or myocardial infarction. Most of the clinical assessment is done qualitatively by visual assessment [38]. Quantitatively, a LV strain indicates deformation of the ventricles, which is used as a parameter in Doppler echoCG analysis to estimate the extent of ischemia, myocardial infarction, and ventricular dyssynchrony.

The LV End-diastole Volume (LVEDV) and the LV End-systole Volume (LVESV) are used to estimate the LV Stroke Volume (LVSV) - the total ejection rate from the LV in each cycle. The Left Ventricle Ejection Fraction (LVEF) is derived by dividing LVESV by LVEDV. LVEF quantifies the fraction of blood pumped out by the LV in each cycle [37].

One technique to derive the LV region is the application of the Hough transform at the end of diastole and model the region using concentric circles. The region showing the maximum projected intensity near the center of concentric circles is a viable candidate for the LV seed [37]. Other properties of the LV region are homogeneity and high grey level. The segmentation, identification techniques, and quantification techniques for other chambers are similar.

CMR images are acquired using fixed periodic time after the occurrence of R-waveforms – the most prominent waveform in ECG associated with the compression of the LV. Patients with arrhythmia have varying heart-cycle which degrades the image quality. There are challenges because different chambers and walls have similar intensity profile, making contrast-based segmentation difficult [39]. Artifacts, noise, lighting conditions and heterogeneity due to blood-flow also affect the intensities [39][40].

VI. PCG ANALYSIS

PCG analysis is based upon signals derived from systolic and diastolic phases of blood-flow turbulence to estimate the occurrence of valvular diseases like stenosis, regurgitation, atherosclerotic disease, and murmur (see Figure 4) [23][41]. However, unlike image-based techniques, PCG cannot accurately localize valvular and muscle-deformity-related diseases, such as myopathy or hypertrophy, and diseases related to blood-leakage.

PCG analysis requires signal analysis in the time-domain using discrete wavelet or packet wavelet transforms [23] or in frequency domain using FFT [27].

Intelligent PCG uses *Time Growing Neural Network* (TGNN) based analysis to segment the time-series of diastolic and systolic sound patterns of stenosis of aortic, pulmonary and tricuspid valves [42]. TGNN combines windowing to extract signal frames with neural network for the classification. Windowing uses a fixed starting point but growing endpoints to identify varying size time-framed windows containing the signal.

A hybrid model combining CNN and LSTM has been used for the classification of PCG [43]. CNN is used to analyze the frequency-related features derived using *Mel Frequency Cepstral Coefficient* (MFCC), and LSTM is used to derive temporal dependencies. Another research combined CNN and bidirectional LSTM to detect aortic

stenosis, mitral stenosis, mitral regurgitation, and mitral valve pro-lapse [44].

VII. MAJOR DATASETS

EchoNet-Dynamic database contains 10,030 labeled apical-4-chamber echocardiography videos and human expert annotations in the form of measurements, tracings, and calculations to study cardiac motion and chamber sizes [45][46]. Data is accompanied by EF values of LV and frame numbers of end-systole and end-diastole frames determined by medical practitioners.

CREATIS repository provides multimodal 2D and 3D cardiac imaging data, application-software, and diagnostics to evaluate computational methods and enhance collaboration around heart imaging and analysis [47]. The repository is a collection of multiple databases: 1) *CAMUS synthetic database* contains 2D apical myocardial motion in four chambers; 2) *Duplex database* contains twenty simulated sequences; 3) *Cetus database* contains 3D echoCG sequences of 45 clinical patients; 4) *Multimodal Straus database* contains 3D echoCG data, cine-MRI and tagged-MRI data of eighteen virtual patients; 5) *Revolus database* contains 2D echoCG of both simulated and actual sequences; 6) *ACDC database* contains 3D CMR data along with manual contouring to mark LV endocardium and RV endocardium for both diastolic and systolic phases; 7) *Minimalist Immediate Mechanical Intervention (MIMI) dataset* comprises a multicentric randomized trial comparing immediate and delayed stenting in 140 patients treated with *Percutaneous Coronary Intervention* (PCI).

The *Harvard dataverse* comprises cardiac imaging datasets, such as CMR images of 35 patients with mitral regurgitation, 4D-flow echoCG image-sequence, and CMR data from 108 subjects (patients and healthy subjects) [48].

The *Heart database* comprises 3D CMR images of LV with automated segmentations validated by clinicians, tools to compute quantitative measures, and software packages for automated image segmentation [49].

The *EMIDEC database* comprises datasets for classifying normal and pathological cases of 150 MRI exams from different patients for studying LV in the cases of myocardial infarction symptoms [50].

The *Physionet database* contains nine PCG databases and applied deep neural networks to classify heart murmurs [51].

The *Cardiac Atlas Project (CAP) database* contains data of asymptomatic and pathological hearts to facilitate collaborative statistical analysis of regional heart shapes and characterize cardiac function for multiple population groups [52].

VIII. DISCUSSION AND CONCLUSION

This review has described the trend of applying deep learning techniques for cardiac image segmentation needed for detecting cardiac chambers, blood-channels, blood-flow and its quantification, and its various defects, such as heart muscle deformation, plaque formation, calcification and valvular defects. It also describes the combination of frame

segmentation and deep learning networks to classify sounds collected in PCG.

The advantage of deep learning is in identifying feature-maps due to intensity and texture variations using convolution layers, and repeated patterns of encoders and decoders, providing temporality as required in analyzing various phases of blood-flow, sound generation and heart-muscle movements in heart cycles. CMR and echoCG images contribute to the better quantification-based image analysis. Traditional quantification techniques such as blood-flow estimates, augmented with deep learning-based boundary detection and strain detection have significantly improved the diagnosis and the classification of heart-defects.

The drawbacks in image analysis are the presence of noise such as speckles due to blood-flow that hinder segmentation, inability to separate similar defects such as ischemic and infarcted regions, inaccurate detection of boundaries of wall chambers, especially when the valves are open. Another drawback is the absence of a large dataset required to improve the accuracy in deep learning techniques. This problem is being resolved progressively, as described in Section VII. Despite clinical validation of results, a major criticism of a deep learning model is the black box approach with no causality-based explanation.

In the next decade, deep learning techniques, combined with continuously reducing cost of use, will be established as a valuable tool for automated diagnostics and fast and accurate decision support system to identify noninvasively the extent and localization of the muscle, motion and valves related cardiac diseases.

REFERENCES

- [1] World Health Organization Facts Sheet, available at [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)), last accessed April 30, 2022.
- [2] S. S. Virani, A. Alonso, H. J. Aparicio, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway et al., "Heart Disease and Stroke Statistics – 2021 Updates," *Circulation*, vol. 143, No. 8, pp. e254-e742, Feb. 2021, doi: 10.1161/CIR.0000000000000950.
- [3] D. V. D. Linde, E. E. M. Konings, M. A. Slager, M. Witsenburg, W. A. Helbing et al., "Birth Prevalence of Congenital Heart Disease Worldwide: A Systematics Review and Meta-analysis," *Journal of American College of Cardiology*, vol. 58, issue 21, pp. 2241-2247, Nov. 2011, doi: 10.1016/j.jacc.2011.08.025.
- [4] K. Maganti, V. H. Rigolin, M. E. Sarano, and R. O. Bonow, "Valvular Heart Disease: Diagnosis and Management," *Symposium on Cardiovascular Diseases, Mayo Clinic Proceedings*, vol. 85, issue 5, pp. 483-500, May 2010, doi: 10.4065/mcp.2009.0706.
- [5] S. J. Al'Aref, K. Anchouche, G. Singh, P. J. Slomka, K. K. Kollu, A. Kumar et al., "Clinical Applications of Machine Learning in Cardiovascular Disease and its Relevance to Cardiac Imaging," *European Heart Journal*, vol. 40, pp. 1975-1986, 2019, doi: 10.1093/eurheartj/ehy404.
- [6] B. Xu, D. Kocyigit, R. Grimm, B. P. Griffin, and F. Cheng, "Applications of Artificial Intelligence in Multimodal Cardiovascular Imaging: A State of Art Review," *Progress in Cardiovascular Diseases*, vol. 63, pp. 367-376, 2020, doi: 10.1016/j.pcad.2020.03.003.
- [7] A. G. Tilkian and M.B. Conover, *Understanding Heart Sounds and Murmurs: With an Introduction To Lung Sounds*, fourth ed., Saunders: Philadelphia, PA, 2001.
- [8] A. K. Bansal, J. I. Khan, and S. K. Alam, "Introduction to computational health informatics," CRC Press: Newyork, NY, 2020, doi: 10.1201/9781003003564.
- [9] S. G. Myerson, J. Francis, and S. Neubauer, *Cardiovascular magnetic resonance*, Oxford University Press: Oxford, UK:, 2013.
- [10] T. Rogers and R. Waksman, "Role of CMR in TAVR," *JACC: Cardiovascular Imaging*, vol. 9, No. 5, pp. 593-602, May 2016, doi: 10.1016/j.jcmg.2016.01.011.
- [11] A. S. Fahmy et al. "Automated cardiac MR scar quantification in hypertrophic cardiomyopathy using deep convolutional neural networks," *JACC Cardiovasc Imaging*, 2018, vol. 11, pp. 1917–18, doi: 10.1016/j.jcmg.2018.04.030.
- [12] F. C. Cobey, V. Patel, and A. Gosling, "The Emperor has no Clothes: Recognizing the Limits of Current Echocardiographic Technology in Perioperative Quantification of Mitral Regurgitation," *Journal of Cardiothoracic and Vascular Anesthesia*, vol. 31, issue 5, pp. 1692-1694, Oct. 2017, doi: 10.1053/j.jvca.2017.03.012.
- [13] C. Knackstedt, S. C. A. M. Bekkers, G. Schummers, M. Schreckenberg, D. Muraru, I. P. Badano et al., "Fully Automated versus Standard Tracking of Left Ventricular Ejection Fraction and Longitudinal Strain: the FAST-EFs Multicenter Study," *Journal of the American College of Cardiology*, vol. 66, no. 13, pp. 1456-1466, Sept. 2015, doi: 10.1016/j.jacc.2015.07.052.
- [14] S. Russel and P. Norvig, "Artificial Intelligence – A Modern Approach," 4th Edition, Pearson Education: Hoboken, NJ, 2020.
- [15] Y. LeCun Y. Bengio, and G. Hinton, "Deep learning," vol. 521, issue 7553, pp. 436-444, May 2015, doi: 10.1038/nature14539.
- [16] L. Deng and D. Yu, "Deep learning methods and applications," *Foundations and Trends in Signal Processing*, vol. 7, issue 3-4, pp. 197-387, June 2014, doi: 10.1561/20000000039.
- [17] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object Detection with Deep Learning – A Review," *IEEE Transactions of Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212-3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.
- [18] Computer generated cross-section 3D model of heart, Wikimedia Commons, Creative License 4.0, available at https://commons.wikimedia.org/wiki/User:DrJanaOfficial#/media/File:CG_Heart.gif, last accessed May 16, 2022.
- [19] Q. Jhang, Y. Liu, J. Mi, X. Wang, X. Liu, F. Zhao et al., "Automatic Assessment of Mitral Regurgitation Severity Using the Mask R-CNN Algorithm with Color Doppler Echocardiography Images," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article Id: 2602688, pp. 1-10, Sept. 2021, doi: 10.1155/2021/2602688.
- [20] Ekko, Ventricular septal defect, Wikimedia Commons, Public domain, available at https://commons.wikimedia.org/wiki/File:Ventricular_Septal_Defect.jpg, last accessed May 16, 2022.
- [21] J. C. C. Moon and M. danne, A CMR image of a benign tumor in a heart, Wikimedia Commons, available at https://commons.wikimedia.org/wiki/File:Myxoma_CMV.gif, last accessed May 16, 2022.
- [22] Phonogram from normal and abnormal heart, Wikimedia Commons, GNU free documentation, License 1.2, https://commons.wikimedia.org/wiki/File:Phonocardiograms_from_normal_and_abnormal_heart_sounds.png#/media/File:P

- honocardiograms_from_normal_and_abnormal_heart_sounds.svg, last accessed May 16, 2022.
- [23] L. H. Cherif, S.M. Debbal, and F. Bereksi-Reguig, "Choice of the Wavelet Analyzing in the Phonocardiogram Signal Analysis using the Discrete and Packet Wavelet Transforms," *Expert Systems with Applications*, vol. 37, issue 2, pp. 913-918, March 2010, doi: 10.1016/j.eswa.2009.09.036.
- [24] F. Kovács, C. Horváth, A. T. Balogh, and G. Hosszú, "Fetal Phonocardiography – Past and Future Possibilities," *Computer Methods and Programs in Biomedicine*, vol. 104, issue 1, pp. 19-25, Oct. 2011, doi: 10.1016/j.cmpb.2010.10.006.
- [25] J. Wang et al., "Automated Interpretation of Congenital Heart Disease from Multi-view Echocardiograms," *Medical Image Analysis*, vol. 69, Article 101942, pp. 1-12, 2021, doi: 10.1016/j.media.2020.101942.
- [26] An illustration of PDA, Wikimedia Commons, Public domain, Available at https://en.wikipedia.org/wiki/Patent_ductus_arteriosus#/media/File:Patent_ductus_arteriosus.svg, last accessed May 16, 2022.
- [27] Z. Abduh, E. A. Neharyb, M. A. Waheda, and Y. M. Kadaha, "Classification of Heart Sounds using Fractional Fourier Transform based Mel-Frequency Spectral Coefficients and Traditional Qualifiers," *Biomedical Signal Processing and Control*, vol. 57, Article Id 101788, pp. 1-11, 2020, doi: 10.1016/j.bspc.2019.101788.
- [28] Z. Akkus et al., "Artificial Intelligence (AI) Empowered Echocardiography Interpretation – A State of Art Review," *Journal of Clinical Medicine*, vol. 10, issue 7, Article Id 1391, pp. 1-16, doi: 10.3390/jcm1007139.
- [29] L. Xu, M. Liu, Z. Shen, H. Wang, X. Liu, and X. Wang, "DW-net a Cascaded Convolution Network for Apical Four-chamber View Segmentation in Fetal Echocardiography," *Computerized Medical Imaging and Graphics*, vol. 80, Article 101690, pp. 1-11, March 2020, doi: 10.1016/j.compmedimag.2019.101690.
- [30] R. Nedadur, B. Wang, and W. Tsang, "Artificial Intelligence for the Echocardiographic Assessment of Valvular Heart Disease," *Heart*, Article Id 319725, pp. 1-8, Feb. 2022, doi: 10.1136/heartjnl-2021-319725.
- [31] I. Wahlang et al., "Deep Learning Methods for Classification of Certain Abnormalities in Echocardiography," *Electronics*, vol. 10, Article 495, pp. 1-20, Feb. 2021, doi: 10.3390/electronics10040495.
- [32] A. K. Pinjari, P. V. Sridevi, and M. N. GiriPrasad, "Segmentation Methods for Severity Regurgitation: A Comparative Analysis," *Research Journal of Recent Sciences*, vol. 3, no. 6, pp. 83-89, June 2014, available at www.isca.in/rjrs/archive/v3/i6/13.ISCA-RJRS-2014-833.pdf.
- [33] A. Ghorbani et al., "Deep Learning Interpretations of Echocardiograms," *NPJ Digital Medicine*, vol. 3, Article 10, pp. 1-10, Jan. 2020, doi: 10.1038/s41746-019-0216-8.
- [34] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint, arXiv:1803.01271*, pp. 1-14, April 2018, doi: 10.48550/arXiv.1803.01271.
- [35] L. Xu et al. "DW-Net: A Cascaded Convolutional Neural Network for Apical Four-chamber View Segmentation in Fetal Echocardiography," *Computerized Medical Imaging and Graphics*, vol. 80, Article Id. 101690, pp. 1-11, March 2020, doi: 10.1016/j.compmedimag.2019.101690.
- [36] C. Constantinidès, E. Roullot, M. Lefort, and F. Frouin, "Fully Automated Segmentation of the Left Ventricle Applied to Cine MR Images: Description and Results on a Database of 45 Subjects," 34th Annual International Conference of the IEEE EMBS, San Diego, CA, USA, Aug.-Sept. 2012, pp. 3207-3210.
- [37] P. Peng, K. Lekadir, A. Gooya, L. Shao, S. E. Petersen, and A. F. Frangi, "A Review of Heart Chamber Segmentation for Structural and Functional Analysis using Cardiac Magnetic Resonance Imaging," *Magn Reson Mater Physics*, vol. 29, issue 2, pp. 155-195, April 2016, doi: 10.1007/s10334-015-0521-4.
- [38] T. Emrich, M. Halfmann, U. J. Schoepf, and K.-F. Kreitner, "CMR for Myocardial Characterization in Ischemic Heart Disease: State of Art and Future Developments," *European Radiology Experimental*, vol. 5, issue 14, pp. 1-13, doi: 10.1186/s41747-021-00208-2.
- [39] L. V. Romaguera, F. P. Romero, C. F. C. Filho, and M. G. Fernandes, "Myocardial Segmentation in Cardiac Magnetic Resonance Images using Fully Convolved Neural Networks," *Biomedical Signal Processing and Control*, vol. 44, pp. 48-57, July 2018, doi: 10.1016/j.bspc.2018.04.008.
- [40] M.R. Avendi, A. Kheradvar, and H. Jafarkhani, "A Combined Deep-learning and Deformable-model Approach to Fully Automatic Segmentation of the Left Ventricle in Cardiac MRI," *Medical Image Analysis*, vol. 30, pp. 108-119, May 2016, doi: 10.1016/j.media.2016.01.005.
- [41] Z. Tariq, S. K. Shah, and Y. Lee, "Automatic Multimodal Heart Disease Classification using Phonocardiogram Signals," *IEEE International Conference on Big Data, Virtual*, Dec. 2020, pp. 3514-3521.
- [42] A. Gharehbaghi, M. Lindén, and A. Babic, "An Artificial Intelligent based Model for Detecting Systolic Pathological Patterns of Phonocardiogram based on Time-growing Neural Network," *Applied Soft Computing Journal*, vol. 83, Article Id 105615, pp. 1-12, Oct. 2019, doi: 10.1016/j.asoc.2019.105615.
- [43] A. N. Netto and L. Abraham, "Detection and Classification of Cardiovascular Disease from Phonocardiogram using Deep Learning Models," *Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2021, pp. 1646-1651, doi: 10.1109/ICESC51422.2021.9532766.
- [44] M. Alkhodari and L. Fraiwan, "Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings," *Computer Methods and Programs in Biomedicine*, vol. 200, Article Id. 105940, pp. 1-25, March 2021, doi: 10.1016/j.cmpb.2021.105940.
- [45] Echonet Dynamic Database. [Online]. Available from: <https://echonet.github.io/dynamic/index.html>.
- [46] D. Ouyang et al., "Video-based AI for beat-to-beat assessment of cardiac function." *Nature*, vol. 580, no. 7802, pp. 252-256, March 2020, doi: 10.1038/s41586-020-2145-8.
- [47] CREATIS – Human Heart Project Databases. [Online]. Available from: <http://humanheart-project.creatis.insa-lyon.fr/databases.html>.
- [48] A. Fahmy et al. Harvard dataverse. [Online]. Available from: <https://dataverse.harvard.edu/dataverse/cardiaccmr>.
- [49] J. Cousty et al., The 4D Heart database, [Online]. Available from: <https://www.laurentnajman.org/heart/index.html>.
- [50] EMIDEC- Automated Evaluation of Myocardial Infarction from Delayed-Enhancement Cardiac MRI database. [Online]. Available from: <http://emidec.com/>.
- [51] Physionet database. [Online]. Available from: <http://www.physionet.org>.
- [52] C. G. Fonseca et al., "The Cardiac Atlas Project—an imaging database for computational modeling and statistical atlases of the heart," *Bioinformatics*, vol. 27, no. 16, pp. 2288-2295, Aug. 2011, doi: 10.1093/bioinformatics/btr360.

Steps towards the Modeling of Animal Vibrissa Modes

Using Adaptive Control

Carsten Behn

Schmalkalden University of Applied Sciences
Schmalkalden, Germany
e-mail: c.behn@hs-sm.de

Moritz Scharff

Wilhelm-Franke-Str. 31
Dresden, Germany
e-mail: moritz.scharff@live.de

Lukas Merker

Technische Universität Ilmenau
Ilmenau, Germany
e-mail: lukas.merker@tu.ilmenau.de

Abstract—The reception of vibrations is a special sense of touch, important for many insects and vertebrates. The latter realize this reception by means of hair-shaped vibrissae all over the body, but especially arranged in the mystacial pad around the snout. Latest research activities focus on modeling these sensor components to explain some biological behaviors/features (Technical Biology) or to investigate their usage for technical applications for, e.g., object surface and/or shape detection (Bionics). In contrast to these works, we focus on the modeling of the dynamic operation modes to be applied for dynamic scanning patterns of objects. We set up a principal mechanical model of a single vibrissa to describe these modes of operation: passive and active vibrissae. They are either used passively to sense environmental forces, e.g., wind, or actively, when they are rhythmically moved to scan objects or surfaces. Consequently, the vibrissa inspired model has to allow for stabilizing and tracking control as well, but yet being able to detect (superimposed) solitary excitations. Hence, the biological paradigm exhibits some adaptive behavior, and so must the controller: to be adaptive in view of both the randomness of the external signals to be suppressed and the uncertainty of the system data.

Keywords—Vibrissa; tactile sensor; mode of operation; adaptive control.

I. INTRODUCTION

Tactile sensor technology seems nowhere near from playing a key role in environmental exploration in mobile robotics, especially when compared to optical sensing technology [1]. Taking a glance at the world of biology, the sense of touch provides necessary information for several species, e.g., mammals, especially nocturnal animals like rats. They exhibit a sophisticated tactile sense organ supporting tactile exploration which complements the visual and aural sensing: the mystacial vibrissae in the snout region, see Figure 1.



Figure 1. Mammals with vibrissae: a cat named 'Alfred' with whiskers (left), and a rat (right) [2].

The rats benefit from their vibrissae in different ways: they

are involved in their social behavior, acquisition of food, locomotion, navigation [3][4], detecting air flows [5]. They enable the rats to detect, localize and recognize objects near their faces [6], including detection of object features (orientation, shape, texture) [7][8].

The mystacial vibrissae are arranged in an array of columns and rows around the snout, see Figure 2.

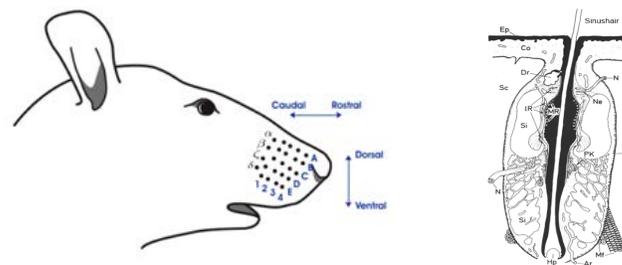


Figure 2. Schematic drawing of the mystacial pad (left) [9], follicle-sinus-complex of a vibrissa (right) [10].

The vibrissa itself (made of dead material) is mainly used as a lever for the force transmission. But, in contrast to ordinary hairs, vibrissae are stiffer and have a hollow conical shape [9]. Each vibrissa in the mystacial pad is embedded in and supported by its own follicle-sinus-complex (FSC), see Figure 2, right. The FSC is characterized by its exceptional arrangement of blood vessels, neural bonding and muscles.

Despite the distribution and difference of the vibrissae in a mystacial pad with respect to their (varying) material and geometrical features (multi-layer system, slenderness, conical shape with inherent curvature inside this pad), we focus on a single vibrissa, its surrounding tissue, its functionality and mode of operation. For a more detailed overview see [11]. Due to [12] the surrounding musculature is divided in extrinsic and intrinsic one. The vibrissa can either be moved passively (e.g., by wind), or actively through alternate contractions of the intrinsic and extrinsic muscles. The resulting rhythmic pro- and retractions of the vibrissa, called whisking, are used to scan surfaces and objects [10]. Relevant information can be perceived by adjusting the frequency and amplitude of the oscillation to each task. By observing rats accomplishing several types of exercises, Berg and Kleinfeld in [13] could distinguish two main *whisking patterns*: *exploratory whisking* when rodents explore their environment with large amplitude sweeps in a low frequency (5-15Hz) range, and *foveal whisking*

when rodents palpate object surfaces with small amplitude, high frequency (15-25Hz) movements. Summarizing, the complex FSC- and muscle-system enables the rodents to use their vibrissae in two different ways (modes of operation):

- In the **passive mode**, the vibrissae are being deflected by external forces (e.g., wind). They return to their rest position passively — thus without any muscle activation.
- In the **active mode**, the vibrissae are swung back-and forward by alternate contractions of the intrinsic and extrinsic muscles. By adjusting the frequency and amplitude of the oscillations, the rodents are able to investigate object surfaces and shapes amazingly fast and with high precision [14].

But, how the animals convert these multiple contacts with single objects into coherent information about their surroundings remains unclear and is not of focus in this work. However, from the point of view from control theory, every biological sensory system has the ability to constantly adapt its sensitivity to its current environment in a way that empowers it to distinguish the relevant information out of the multitude of negligible stimuli (\rightarrow *adaptive system*). Several *control strategies* enhance the relevance-oriented stimulus processing:

- a feedback-loop (closed-loop control system) enables the rodents to immediately react to an object contact: they slow down the concerned vibrissae, diminishing the occurring wear-out effect on the hair [15][16],
- depending on the mode (passive or active) and the expectations of the rodent, the neurons reaction is being suppressed, enhanced or left unaltered [17]–[19].

Therefore, this biological sensor system is highly interesting for applications in the field of autonomous robotics, since tactile sensors can offer reliable information, where conventional sensors fail (in dark, smoky or noisy environments).

The remainder of the paper at hand is structured as follows: In Section II, we give a very brief overview on mathematical models to describe the rigid-body motion of vibrissa-like mechanical systems. Then, in Section III, we present the vibrissa-inspired sensor model, derive the equations of motion, and present some aspects of the mathematical treatment and the modes of operation in Section IV. Section V is devoted to numerical simulations. Finally, the results of the present paper are summed up and some future research subjects are identified in Section VI.

II. STATE OF THE ART - MODELING ANIMAL VIBRISSAE

An intensive literature overview of technical vibrissa models (rigid body and continuum) has been presented in [20]–[22]. Here, we briefly discuss some models therein, where we restrict the investigations to rigid body models, continuous models are in recent exploration.

The following summarizes the relevant information of the encountered mechanical models found in the literature for this paper.

- Mitchinson in 2004 [23] and 2007 [16] - *Model of the FSC*
 - ⊖ too complex for a technical implementation
 - ⊕ Determination of spring and damping coefficients for the FSC
- Berg in 2003 [13] and Hill in 2008 [24] - *Model of the musculature in the mystacial pad*
 - ⊕ Implementation of intrinsic and extrinsic musculature
 - ⊕ Simulating the viscoelastic properties of the skin
 - \hookrightarrow Determination of spring and damping coefficients for the skin
 - ⊖ Negligence of the viscoelastic properties of the FSC
 - ⊖ Connection between the follicles
 - \hookrightarrow leads to complex control strategy and high control effort
- Berg in 2003 [13] and Hill in 2008 [24] - *Determination of the range of movement of the vibrissa*
 - $\hookrightarrow \varphi_{Rest} \approx 80^\circ$
hence: angular deflection in rostral direction $\varphi_{rest} + 65^\circ$, angular deflection in caudal direction $\varphi_{rest} - 35^\circ$, translatory shift $\approx 5mm$ in caudal-rostral direction, translatory shift $\approx 3mm$ in dorsal-ventral direction
- models by Simony [25] and Haiderliu [26] in 2010 are too complex in their structure and neglected for discussion here
- Behn in 2013 [21] - *Model of stimulus transmission*
 - ⊕ Implementation of viscoelastic properties of FSC
 - ⊖ Negligence of the viscoelastic properties of the skin

The goal is *not* to recreate an exact copy of the biological system, but *to implement the specific characteristics of the vibrissa needed for the detection of useful information in challenging surroundings in a mechanical model*. Principally, the tenor of our investigations is from bionics: modeling live paradigms, exploiting corresponding mathematical models in order to understand details of internal processes and, possibly, coming to artificial prototypes (e.g., sensors in robotics). We point out that we focus on a single vibrissa and not on a tuft of various vibrissae as in [26].

III. A FIRST MODEL OF A SINGLE VIBRISSA

Following [12] and as mentioned in Section I, the vibrissa is supported by its FSC. Important parts of it are an enveloping chamber with controllable blood supply, and intrinsic and extrinsic muscles which control the vibrissa motions. This biological description suggests physical models as sketched in Figure 3, and these then led us to the mathematical model based on the pendulum device in Figure 4. The model has a viscoelastic support (skin level) and is acted upon by a force excitation $F(\cdot)$ and a control torque $M_u(\cdot)$. The control torque is due to control impact by the musculature [12]. The equations

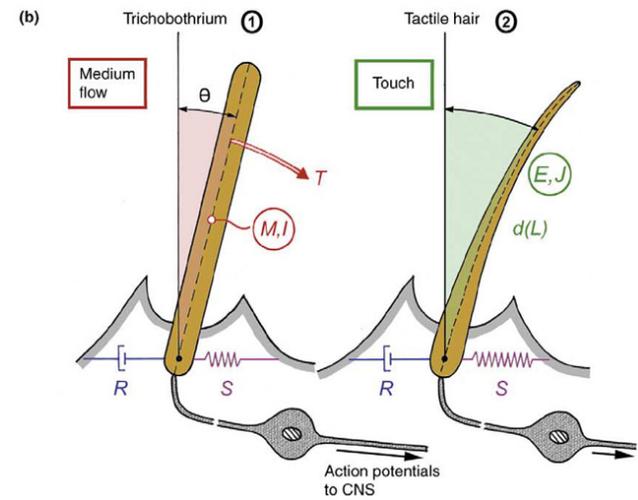


Figure 3. Scheme of a vibrissa. [27].

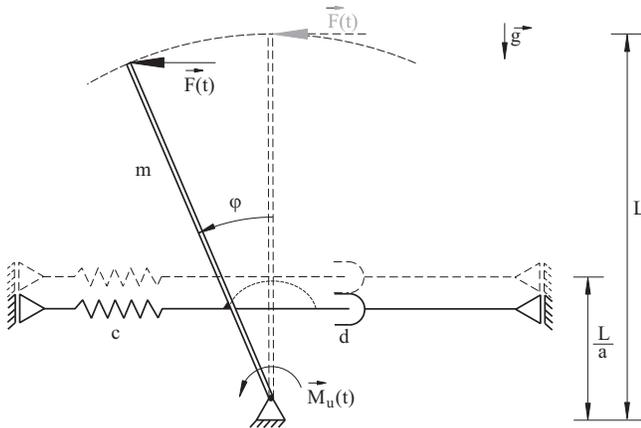


Figure 4. A mechanical model of a vibrissa.

of motion are given by the Principle of Angular Momentum:

$$\begin{aligned}
 J_0 \ddot{\varphi}(t) &= \frac{L}{2} m g \sin(\varphi(t)) \\
 &- \frac{L^2}{a^2} c \sin(\varphi(t)) \cos(\varphi(t)) - \frac{L^2}{a^2} d \cos(\varphi(t))^2 \dot{\varphi}(t) \\
 &+ L \cos(\varphi(t)) F(t) + M_u(t), \quad (1)
 \end{aligned}$$

where $J_0 := \frac{1}{3} m L^2$ is the moment of inertia about the pivot. The output of the system is $y(t) = \varphi(t)$.

IV. MATHEMATICAL BACKGROUND

Dealing with sophisticated, biologically inspired systems, one cannot expect to have all information about it. Instead only structural properties are known. In doing so, we assume that internal data (e.g., mass, spring stiffness, and damping ratio) are unknown. Furthermore, with respect to a changing and therefore uncertain environment, we assume the external excitations to be unknown which leads us to the treatment of *highly uncertain (control) system* of known structure given in (1). Note that this system belongs to the system class presented

in [21]: finite-dimensional, nonlinearly perturbed, m -input $u(\cdot)$, m -output $y(\cdot)$ systems of relative degree two with known sign of the high-frequency gain spectrum $(G) \in \mathbb{C}_+$:

$$\dot{x} = f(t, x, \dot{x}) + G u, \quad y = C x. \quad (2)$$

To dominate this system with uncertain parameters, we have to design an adaptive controller which learns from the behavior of the system in automatically adjusting its (gain) parameters to achieve a desired control objective. Because the system is excited by unknown forces (e.g., possible wind from the environment), we try to design a universal feedback controller which adaptively compensates this unknown excitation and leaves the system in a desired operation pattern.

Since we are dealing with an uncertain, randomly perturbed, non-autonomous system, particular attention is paid to the adaptive λ -tracking control objective: determine a universal λ -controller, which learns from the behavior of the system and automatically adjusts its parameters such that the system tracks a given reference signal $y_{\text{ref}}(\cdot)$ (representing a desired mode of operation) with a prescribed accuracy λ . $\lambda > 0$ denotes the size of the feasible tracking error, which means that the error $e(t) := y(t) - y_{\text{ref}}(t)$ is forced, via the adaptive feedback mechanism, towards a ball around zero of arbitrary small pre-specified radius λ [28]. Choosing $y_{\text{ref}}(\cdot) \equiv 0$, $\lambda = 0$ we arrive at the so-called *adaptive stabilization control objective*.

A preferred control strategy is the following [28]:

$$\left. \begin{aligned}
 e(t) &:= y(t) - y_{\text{ref}}(t), \\
 u(t) &= -\left(k(t)e(t) + \kappa \frac{d}{dt}(k(t)e(t))\right), \\
 \dot{k}(t) &= \gamma \left(\max\{0, \|e(t)\| - \lambda\}\right)^2,
 \end{aligned} \right\} \quad (3)$$

with $k(0) = k_0 \in \mathbb{R}$, $\lambda > 0$, $\kappa = 1$ (just guaranteeing dimensions) and $\gamma \gg 1$.

This controller consists of a very simple feedback mechanism and adaptation law, it is only based on the output of the system and its time derivative - no knowledge of the system parameters is required.

As mentioned in Section I, the biological paradigm vibrissa exhibits two basic modes of operation: a passive and an active mode. Expressing these modes using the presented adaptive control strategies, we conclude that a vibrissa in *passive mode* is a system to be *stabilized* under permanent excitation while enabling to *detect* external extra-perturbations. In an *active mode*, it is to *track* a given oscillatory motion pattern in order to enable the system to recognize, e.g., the surface texture of an external object contacts. Therefore, we can identify these two basic operation modes with

- passive mode: λ -stabilization (see Figure 5),
- active mode: λ -tracking (see Figure 6).

Note: Because of the randomness of external signals and unknown system parameters, we have to choose the presented adaptive control strategy. The given output of system (1) is $y(t) = \varphi(t)$, hence the equations belong to (2) and the presented controllers achieve their objectives. Now, we apply controller (3) (obviously replacing $u(\cdot)$ by $M_u(\cdot)$) to achieve λ -stabilization ($y_{\text{ref}}(\cdot) \equiv 0$ in passive mode) and λ -tracking (in active mode).

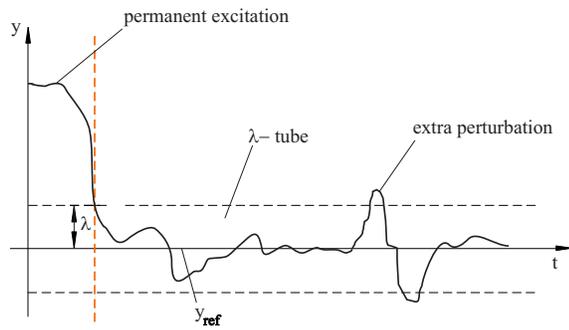


Figure 5. Vibrissa deflection in passive mode.

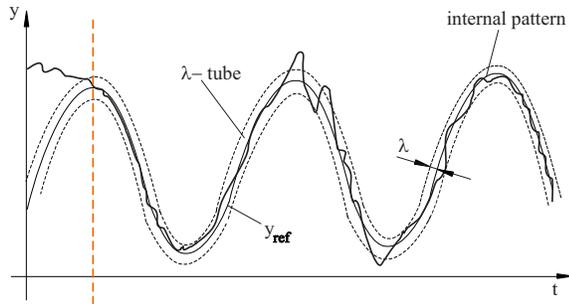
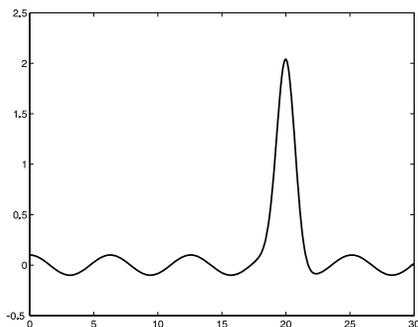


Figure 6. Vibrissa deflection in active mode.

V. SIMULATIONS

We point out that the adaptive nature of the controller is expressed by the *arbitrary choice* of the system parameters. Obviously numerical simulation needs fixed (and known) system data, but the controller *adjusts* its gain parameter *to each set* of system data. Hence, we choose the following parameters (for all simulations, further data will be given on the spot), see [24]:

- vibrissa: $m = 0.003 \text{ kg}$, $c = 5.7 \frac{\text{N}}{\text{m}}$, $d = 0.2 \frac{\text{Ns}}{\text{m}}$, $L = 0.04 \text{ m}$, $a = \frac{L}{10} = 0.004 \text{ m}$, $(\varphi(0), \dot{\varphi}(0)) = (0 \text{ rad}, 0 \frac{\text{rad}}{\text{s}})$;
- environment: $t \mapsto F(t) = (0.1 \cos(t) + 2e^{-(t-20)^2}) \text{ N}$ (see Figure 7) which represents a (small) permanent oscillation with a gust of wind; and $g = 9.81 \frac{\text{m}}{\text{s}^2}$;


 Figure 7. Excitation, F vs. t .

- λ -tracker: $\lambda = 0.2^\circ \approx 0.064 \pi$, $\gamma = 50$ (additional

parameter to increase the convergence of the controller gain [28]) and $k_0 = 0$.

The following reference signals are used for operation in different modes:

- In the passive mode, the rod motion only needs to be stabilized, therefore it can be simulated with the reference signal (4).

$$t \mapsto \varphi_{ref0}(t) = 0 \quad (4)$$

- The active mode has to be implemented with reference signals performing periodical oscillations enabling the rod to either explore its surroundings or to scan specific objects – and since rodents use two different kinds of oscillations depending on the task, the exploratory and foveal whisking are simulated by two different reference signals. Rodents employ large amplitude sweeps in a low frequency range (5 – 15 Hz) to investigate their environment. As the range of movement of the biological vibrissa amounts to ca. $100^\circ = 1.74 \text{ rad}$, the amplitude of the exploratory reference signal (5) can be chosen to $A = \frac{1.74}{2} \approx 0.8 \text{ rad}$. The frequency of the signal $\varphi_{ref1}(t)$ has been set to $f = 5 \text{ Hz}$ according to the findings in [13]. Foveal whisking has been implemented with the reference signal (6), using an amplitude of $A = 0.2 \text{ rad} \approx 12^\circ$ and a frequency of $f = 25 \text{ Hz}$, since rodents scan specific objects with small amplitude, high frequency movements (15 – 25 Hz).

$$t \mapsto \varphi_{ref1}(t) = 0.8 \sin(2\pi 5 t) \quad (5)$$

$$t \mapsto \varphi_{ref2}(t) = 0.2 \sin(2\pi 25 t) \quad (6)$$

A. Passive mode: λ -stabilization of $\varphi_{ref0}(\cdot)$

We use controller (3) for stabilization ($y_{ref}(\cdot) \equiv 0$, rest position to be tracked).

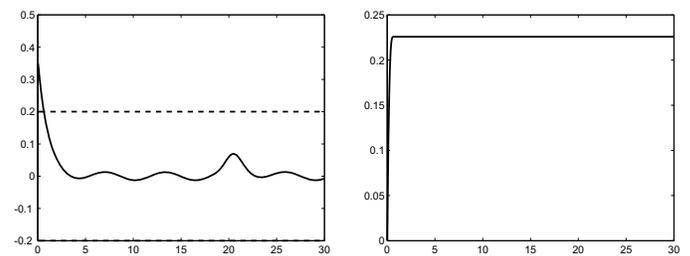
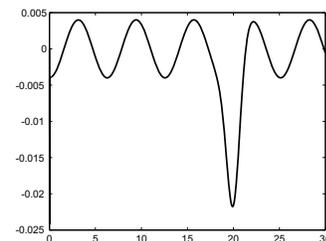

 Figure 8. Output and λ -strip (left), gain parameter (right), all vs. t .

 Figure 9. Control torque M_u vs. t .

Figure 8 (left) shows a good stabilization - the output is captured by the tube. Figure 8 (right) shows the convergence of the gain parameter to a constant value, and Figure 9 the necessary control torque, which reflects the local disturbance around $t = 20$.

B. Active mode 1: λ -tracking of $\varphi_{ref1}(\cdot)$

Now we pass to an active mode. Again we use controller (3) to track reference signal $\varphi_{ref1}(\cdot)$.

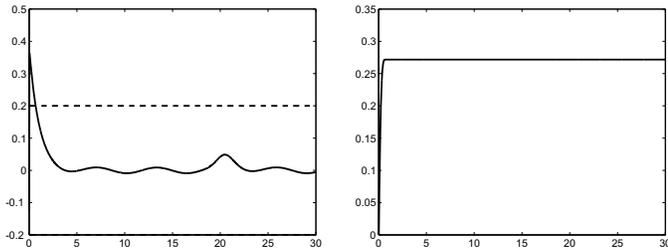


Figure 10. Error e with λ -strip (left), and gain parameter k (right), all vs. t .

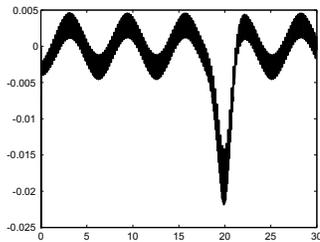


Figure 11. Control torque M_u , all vs. t .

We have a good tracking behavior, the output is captured by the tube. This seems to be true because the error is captured by the tube around zero, see Figure 10 (left). Let us point out that we omit the figure of the output for reasons of presentation (oscillations with high frequency to be tracked) and focus on the error in the following. The controller gain is given in Figure 10 (right). Figure 11 shows the control torque that very weakly responds to the local disturbance.

C. Active mode 2: λ -tracking of $\varphi_{ref2}(\cdot)$

Here, we simulate another active mode in tracking reference signal $\varphi_{ref2}(\cdot)$ using controller (3).

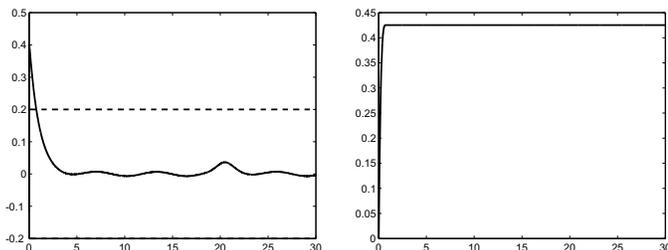


Figure 12. Error e with λ -strip (left), and gain parameter k (right), all vs. t .

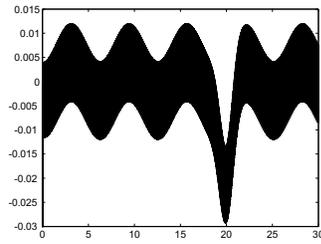


Figure 13. Control torque M_u vs. t .

Again, and also in simulation of another active mode (foveal whisking) we have a good tracking behavior because the controller works effectively, see Figures 12 and 13. But, it is quite hard to detect the gust of wind in the system variable. Only the control torque reflects the “peak” around $t = 20$ s, see Figure 13. Therefore, we try to detect the additional excitation by means of other observables.

D. Improved adaptation laws

The foregoing simulation shows that the system is not really sensitive to notice the excitation peak around $t = 20$, though tracking and stabilization are essentially guaranteed. Moreover, after the control objective is achieved the gain parameter still stays at its high value. Figures 8 (right), 10 (right) and 12 (right) show the monotonic increase of $k(\cdot)$ towards a limit k_∞ . But if some perturbation repeatedly caused the output to leave the λ -strip then $k(t)$ would take larger values again and again. The aim is now to avoid this drawback.

We now apply an improved adaptation law, see [29], that makes $k(\cdot)$ decrease as long as further high level is not necessary. We distinguish three cases: 1. increasing $k(\cdot)$ while e is outside the tube, 2. constant $k(\cdot)$ after e entered the tube - no longer than a pre-specified duration t_d of stay, and 3. decreasing $k(\cdot)$ after this duration has been exceeded. For instance:

$$\dot{k}(t) = \begin{cases} \gamma (\|e(t)\| - \lambda)^2, & \text{if } \|e(t)\| \geq \lambda, \\ 0, & \text{if } (\|e(t)\| < \lambda) \wedge (t - t_E < t_d), \\ -\sigma k(t), & \text{if } (\|e(t)\| < \lambda) \wedge (t - t_E \geq t_d), \end{cases} \quad (7)$$

with given $\sigma > 0$, $\gamma \gg 1$, and $t_d > 0$, whereas the entry time t_E is an internal variable.

Choosing this new adaptor (7) with $\sigma = 0.2$ (moderate exponential decay rate [21]) and $t_d = 1$ (moderate choice of the duration of stay within the λ -tube [29]), we obtain the following tracking results in an active mode, e.g., λ -tracking of $\varphi_{ref1}(\cdot)$:

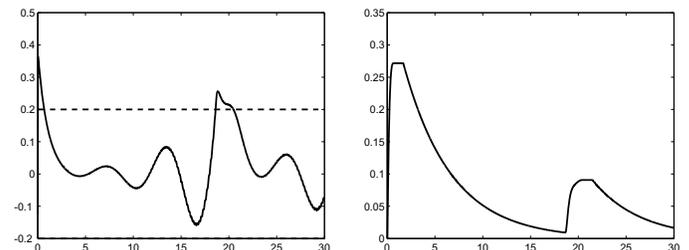
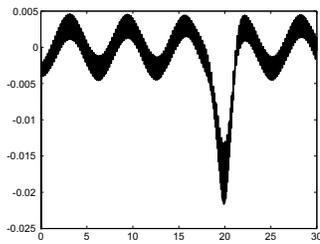


Figure 14. Error e and λ -strip (left), and gain parameter k (right), all vs. t .

Figure 15. Control torque M_u vs. t .

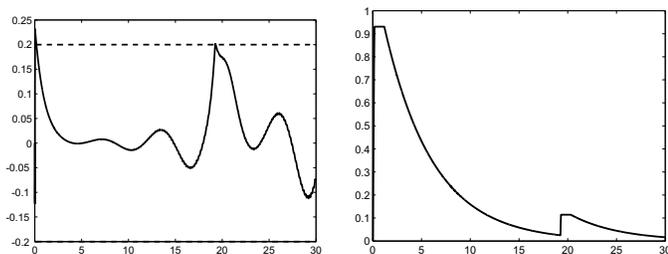
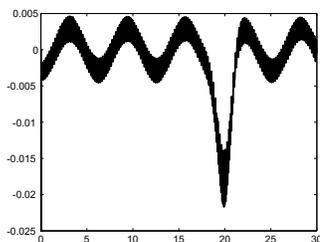
The output is forced into the tube, see Figure 14 (left). The gain $k(\cdot)$ immediately decreases, see Figure 14 (right), after a t_d -stay of e in the tube, see Figure 14 (left). Apparently the decrease is too fast: e leaves the tube again but, as a consequence of \dot{k} being proportional to the *square* of the (small) deviation from the λ -tube, k increases too slowly as to force e quickly back to the tube. The necessary control input is displayed in Figure 15.

In order to make the attraction of the tube stronger, we use different exponents for large/small distance from the tube, see [29]. For instance:

$$\dot{k}(t) = \begin{cases} \gamma \left(\|e(t)\| - \lambda \right)^2, & \text{if } \|e(t)\| \geq \lambda + 1, \\ \gamma \left(\|e(t)\| - \lambda \right)^{0.5}, & \text{if } \lambda + 1 > \|e(t)\| \geq \lambda, \\ 0, & \text{if } \left(\|e(t)\| < \lambda \right) \wedge (t - t_E < t_d), \\ -\sigma k(t), & \text{if } \left(\|e(t)\| < \lambda \right) \wedge (t - t_E \geq t_d), \end{cases}$$

with σ, γ, t_d, t_E as before.

Using this adaptor, we obtain the results from Figures 16 and 17.

Figure 16. Error e and λ -strip (left), and gain parameter k , all vs. t .Figure 17. Control torque M_u vs. t .

Here, one cannot clearly see the advantage of two alternating exponents. The gain parameter $k(\cdot)$ increases very fast in

the beginning of the simulation, see Figure 16 (right). This is due to the “switching on” of the controller. It forces e faster into the tube at cost of a high k . Then, k decreases if e stays in the tube for a duration larger than t_d , see Figure 16 (right). e is forced back into the λ -tube very fast, see Figure 16, because the attraction of the tube is stronger for small deviations $\|e(t)\| - \lambda > 0$. This is on the expense of a tolerable overshooting of k at the beginning, mentioned above. But, interestingly, we are also able to detect solitary excitations in observing the gain $k(\cdot)$ instead of $M_u(\cdot)$, see Figure 17.

VI. CONCLUSIONS

The foregoing considerations have shown that adaptive control is promising in application to vibrissae systems. In particular, it allows, based on the dynamical equations, to describe two main modes of operation of vibrissae: passive and active ones. Improved adaptive controllers are useful and should be developed further.

Current tasks include

- separation of an extra receptor from vibrissa,
- allowing for input and output disturbances,
- identification techniques to get knowledge of solitary excitations.

Near future tasks are

- replace circular pendulum by spherical one (spatial receptivity),
- tuning of the adaptors (i.e., to find favorable values of γ, σ, t_d), and investigating rates of convergence,
- hardware experiments to validate the theory,
- to develop elastic vibrissae models, e.g., multi-body systems and/or continuum systems,
- consider a tuft of vibrissae (modeling the intrinsic musculature), not only a single vibrissa.

After doing this, one can think about the application of the presented tracking scenario to replace the kinematic drive / kinematic scanning trail in present application for recognition of object contours and/or texture in, e.g., [30][31].

REFERENCES

- [1] M. R. Cutkosky, R. D. Howe, and W. R. Provancher, “Force and Tactile Sensors,” in: Springer Handbook of Robotics (Eds. B. Siciliano, O. Khatib), Springer, Heidelberg, pp. 445–476, 2008.
- [2] T. J. Prescott, M. J. Pearson, B. Mitchinson, J. C. W. Sullivan, and A. Pipe, “Whisking with robots: from rat vibrissae to biomimetic technology for active touch,” IEEE Robotics & Automation Magazine, vol. 16, pp. 42–50, 2009.
- [3] A. S. Ahl, “The role of vibrissae in behavior: a status review,” Veterinary Research Communications, vol. 10, pp. 245–268, 1986.
- [4] M. J. Z. Hartmann, “A night in the life of a rat: vibrissal mechanics and tactile exploration,” Annals of the New York Academy of Sciences, vol. 1225, pp. 110–118, 2011.
- [5] Y. S. W. Yu, M. M. Graff, and M. J. Z. Hartmann, “Mechanical responses of rat vibrissae to airflow,” Journal of Experimental Biology, vol. 219, pp. 937–948, 2016.

- [6] S. J. Whiteley, P. M. Knutsen, D. W. Matthews, and D. Kleinfeld, "Deflection of a vibrissa leads to a gradient of strain across mechanoreceptors in a mystacial follicle," *Journal of Neurophysiology*, vol. 114, pp. 138–145, 2015.
- [7] M. Brecht, B. Preilowski, and M. M. Merzenich, "Functional architecture of the mystacial vibrissae," *Behavioural Brain Research*, vol. 84, pp. 81–97, 1997.
- [8] G. E. Carvell and D. J. Simons, "Biometric analyses of vibrissal tactile discrimination in the rat," *Journal of Neuroscience*, vol. 10, pp. 2638–2648, 1990.
- [9] D. Voges et al., "Structural characterization of the whisker system of the rat," *IEEE Sensors Journal*, vol. 12(2), pp. 332–339, 2012.
- [10] A. Schierloh, "Neuronal networks and their plasticity within the barrel cortex of a rat," PhD-Thesis (in german), Technische Universität München, Germany, 2003.
- [11] M. Scharff, "Bio-inspired tactile sensing – Analysis of the inherent characteristics of a vibrissa-like tactile sensor," PhD-Thesis, Technische Universität Ilmenau, Germany, 2021.
- [12] J. Dörfel, "The musculature of the mystacial vibrissae of the white mouse," *Journal of Anatomy*, vol. 135, pp. 147–154, 1982.
- [13] R. W. Berg and D. Kleinfeld, "Rhythmic whisking by rat: retraction as well as protraction of the vibrissae is under active muscular control," *Journal of Neurophysiology*, vol. 89, pp. 104–117, 2003.
- [14] M. J. Z. Hartmann and J. H. Solomon, "Robotic whiskers used to sense features: Whiskers mimicking those of seals or rats might be useful for underwater tracking or tactile exploration," *NATURE*, vol. 443, p. 525, 2006.
- [15] B. Mitchinson, M. Pearson, C. Melhuish, and T. J. Prescott, "A model of sensorimotor coordination in the rat whisker system," in: *From animals to animats 9: Proc. 9th. Int. Conf. on Simulation of Adaptive Behaviour*, pp. 77–88, 2006.
- [16] B. Mitchinson, C. J. Martin, R. A. Grant, and T. J. Prescott, "Feedback control in active sensing: rat exploratory whisking is modulated by environmental contact," *Proc. R. Soc. B*, vol. 274, pp. 1035–1041, 2007.
- [17] D. Derdikman et al., "Layer-specific touch-dependent facilitation and depression in the somatosensory cortex during active whisking," *The Journal of Neuroscience*, vol. 26(37), pp. 9358–9547, 2006.
- [18] J. C. Curtis and D. Kleinfeld, "Seeing what the mouse sees with its vibrissae: a matter of behavioural state," *Neuron*, vol. 50(4), pp. 524–526, 2006.
- [19] E. E. Fanselow and M. A. L. Nicolelis, "Behavioral modulation of tactile responses in the rat somatosensory system," *The Journal of Neuroscience*, vol. 19(17), pp. 7603–7616, 1999.
- [20] T. A. Schmitz, "Development and analysis of biologically inspired sensor systems with higher degrees of freedom using the example vibrissa," Master thesis (in german), Technische Universität Ilmenau, Germany, 2011.
- [21] C. Behn, "Mathematical modeling and control of biologically inspired uncertain motion systems with adaptive features," Habilitation thesis, Technische Universität Ilmenau, Germany, 2013.
- [22] F. A. Lucianna, A. L. Albarracín, S. M. Vrech, F. D. Farfán, and C. J. Felice, "The mathematical whisker: a review of numerical models of rat's vibrissa biomechanics," *Journal of Biomechanics*, vol. 49, pp. 2007–2014, 2016.
- [23] B. Mitchinson et al., "Empirically inspired simulated electro-mechanical model of the rat mystacial follicle-sinus complex," *Proc. R. Soc. Lond.*, vol. 271, pp. 2509–2516, 2004.
- [24] D. Hill, R. Bermejo, P. Zeigler, and D. Kleinfeld, "Biomechanics of the vibrissa motor plant in rat: rhythmic whisking consists of triphasic neuromuscular activity," *Journal of Neuroscience*, vol. 28, pp. 3438–3455, 2008.
- [25] E. Simony et al., "Temporal and spatial characteristics of vibrissa responses to motor commands," *Journal of Neuroscience*, vol. 30, pp. 8935–8952, 2010.
- [26] S. Haiderliu, E. Simony, D. Golomb, and E. Ahissar, "Muscle architecture in the mystacial pad of the rat," *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, vol. 293, pp. 1192–1206, 2010.
- [27] F. G. Barth, "Spider mechanoreceptors," *Current Opinion in Neurobiology*, vol. 14, pp. 415–422, 2004.
- [28] C. Behn, "Adaptive control of straight worms without derivative measurement," *Multibody System Dynamics*, vol. 26, pp. 213–243, 2011.
- [29] C. Behn, "Modeling, analysis and control of mechanoreceptors with adaptive features," *Informatics in Control, Automation and Robotics – Lecture Notes in Electrical Engineering (LNEE)*, vol. 325 (Editors J.-L. Ferrier et al.), Springer International Publishing Switzerland, pp. 349–366, 2014.
- [30] M. Scharff, J. H. Alencastre, and C. Behn, "Detection of surface texture with an artificial tactile sensor," *Interdisciplinary Applications of Kinematics – Mechanisms and Machine Science*, vol. 71 (Eds. A. Kecskeméthy et al.) Springer, Cham, pp. 43–50, 2019.
- [31] L. Merker, M. Scharff, K. Zimmermann, and C. Behn, "Surface Sensing of 3D Objects Using Vibrissa-like Intelligent Tactile Sensors," *Proceedings INTELLI 2020: The Ninth International Conference on Intelligent Systems and Applications, Porto (Portugal), IARIA*, pp. 18–23, 2020.