ICMV 2017
Vienna, Austria

ICMV 2017

CONFERENCE ABSTRATCS

Nov. 13-15, 2017 | Vienna, Austria

2017 The 10th International Conference on Machine Vision

Fleming´s Conference Hotel Wien

Address: Neubaugürtel 26-28, 1070 Wien, Austria

E-mail: wien@flemings-hotels.com / conference.vie@flemings-hotels.com.
Website: https://www.flemings-hotels.com/en/hotels-residences/vienna/flemings-conference-hotel-wien/hotel-rooms

Table of Contents

Welcome Address····· 2

Organizing Committee····· 3

Local Information····· 5

Instructions for Oral&Poster Presentations····· 7

Instructions of Keynote and Plenary Speakers ····· 9

Daily Schedule of Events····· 15

Technical Program (Oral Presentations)····· 17

Technical Program (Poster Presentations)····· 54

Listeners····· 63

Author Index····· 64

One Day Visit ····· 68

Welcome Address

We are delighted to welcome you to ICMV 2017 conference, which will be held at **Fleming´s Conference Hotel Wien,Vienna, Austria,** from Nov. 13-15, 2017.

After several rounds review procedure, the program committee accepted those papers to be published in conference proceedings. We wish to express our sincere appreciation to all the individulas who have contributed to ICMV 2017 in various ways. Special thanks are extended to our colleagues in program committee for their thorough review of all the submissions, which is vital to the success of the conference, and also to the members in the organizing committee and the volunteers who had delicated their time and efforts in planning, promoting, organizing and helping the conference. Last but not least, our special thanks go to conference chairs Prof. Antanas Verikas, Halmstad University, Sweden, Prof. Petia Radeva, University of Barcelona, Spain, Prof. Dmitry Nikolaev, Institute for Information Transmission Problems, RAS (IITP RAS), Russia, for all the kind and patient support and assistance they offered to our whole conference procedure. Without their support, our conference could not be prepared so smoothly, thanks again.

This conference program is highlighted by five Keynote Speakers: Prof. Antanas Verikas, Halmstad University, Sweden, Prof. Petia Radeva, University of Barcelona, Spain, Prof. Dmitry Nikolaev, Institute for Information Transmission Problems, RAS (IITP RAS), Russia, Prof. Dr. Petra Perner (IAPR Fellow),Institute of Computer Vision and Applied Computer Sciences, Leipzig,Germany,Prof. Alexander Bernstein, Skolkovo Institute of Science and Technology (Skoltech), Moscow, Russia Federation.

One best presentation will be selected from each session, evaluated from: Originality; Applicability; Technical Merit; PPT; English. The best one will be announced at the end of each Session, and awarded the certificate over the Dinner.

Vienna is the capital and largest city of Austria and one of the nine states of Austria. Apart from being regarded as the City of Music because of its musical legacy, Vienna is also said to be "The City of Dreams" because it was home to the world's first psycho-analyst – Sigmund Freud. It is well known for having played an essential role as a leading European music centre, from the great age of Viennese Classicism through the early part of the 20th century. Vienna was the world's number-one destination for international congresses and conventions. It attracts over 6.8 million tourists a year.

We wish you a enjoyable visit in Vienna!

ICMV 2017
Conference Organizing Committee
2017. 10. 25

Organizing Committee

International Advisory Committee

Prof. Petra Perner, Institute of Computer Vision and Applied Computer Sciences, Germany
Prof. Alexander Bernstein, Skolkovo Institute of Science and Technology (Skoltech), Russia Federation

Conference Chairs

Prof. Antanas Verikas, Halmstad University, Sweden
Prof. Petia Radeva, University of Barcelona, Spain
Prof. Dmitry Nikolaev, Institute for Information Transmission Problems, RAS (IITP RAS), Russia

Program Chairs

Prof. Aristidis Likas, University of Ioannina, Greece
Prof. Evgeni Magid, Kazan Federal University, Russia
Prof. Pino Caballero-Gil, University of La Laguna,Spain

Publication Chair

Dr. Jianhong Zhou, Sichuan University Jincheng College, China

Technical Committee

Dr. Hao Ji, California State Polytechnic University, USA
Dr. Gernot Stübl, Visual Computing team, Austria
Assoc. Prof. Zhenyu Liu, Shenyang University Technology, China
Prof. Manjunath V. Joshi, Dhirubhai Ambani Institute of Information and Communication Technology, India
Assoc. Prof. Victoria A. Sablina, RSREU, Russia
Prof. Alexander Privalov,Tula State Pedagogical University, Russia
Assoc. Prof. José Miguel Puerta Callejón, UCLM, Spain
Prof. Laurent Wendling, Paris Descartes University, France
Assoc. Prof. Danping Jia, Shenyang University of Technology, China
Prof. Manjunath V. Joshi, Indian Institute of Information Technology, India
Prof. Fottner Johannes, Technical University of Munich Institute for Materials Handling, Germany
Prof. Laure Tougne, Université de Lyon, France
Assoc. Prof. Yun Gao, Yunnan University, China
Prof. Brian W. Surgenor, Queen's University, Canada
Prof. Mourad Zaied, REGIM, Tunisia
Prof. Sharifah Mumtazah Syed Ahmad, University Putra Malaysia, Malaysia
Prof. KOUAHLA Zineddine, LABSTIC-LINA, France

Prof. Pino Caballero-Gil, University of La Laguna, Spain
Prof. Mircea-Florin Vaida, Technical University of Cluj-Napoca, Romania
Prof. Kamel HAMROUNI, Ecole Nationale d'Ingénieurs de Tunis,Tunisia
Prof. Angelo Cangelosi, University of Plymouth, UK
Prof. Igor Faradjev, Russian Academy of Sciences, Moscow, Russia
Prof. Oleg Slavin, Russian Academy of Sciences, Moscow, Russia
Prof. Frédéric Morain-Nicolier, Université de Reims, France
Prof. Mohamed Hammami, Miracl-FS,Tunisia
Prof. Luís Nunes, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal
Prof. Batouche Mohamed, University of Constantine 2, Algeria
Prof. Deng Huan, Sichuan University, China
Prof. Chokri Ben Ammar, National Engineering School of Sfax, Tunisia
Prof. M. Iqbal b. Saripan, University Putra Malaysia, Malaysia
Prof. Henri NICOLAS, University of Bordeaux, France
Prof. Assia KOURGLI, Faculté d'Electronique et d'Informatique, LTIR, Algérie
Prof. Qing Linbo, Sichuan University, China
Dr. Amir Ghaderi, University of Texas, USA
Dr. Najet Arous, Enit,Tunisia
Prof. Maya Dimitrova, Bulgarian Academy of Sciences, Bulgaria
Prof. Henryk Palus, Silesian University of Technology, Poland
Prof. Evgeny Myasnikov, Samara National Research University, Russia
Prof. Simon Karpenko, Institute for Information Transmission Problems of the Russia Academy of Sciences, Russia
Assoc. Prof. Umair Ali Khan, Quaid-e-Awam University of Engineering, Science & Technology, Pakistan
Prof. Aleksandar Markoski, "St Kliment Ohridski" University – Bitola, Macedonia
Prof. Sei-ichiro Kamata, Waseda University, Japan
Prof. Mohamed El-Sayed Farag, Al-Azhar University, Egypt
Prof. Mehmet CUNKAS, Selcuk University, Turkey
Prof. Mohamed Rizon, Universiti Sultan Zainal Abidin, Malaysia
Prof. Atsuo Yoshitaka, Japan Advanced Institute of Science and Technology, Japan
Dr. Francesco Ciompi, Radboud University, The Netherlands
Dr. Jose M. Massa, UNICEN University, Argentina
Prof. Marcos Orgega, University of A Coruña, Spain
Assoc. Prof. Xose M. Pardo Lopez, University of Santiago de Compostela, Spain
Assoc. Prof. Sadekov Rinat Nailevish, MEI "Institute of Engineering Physics", Russia
Dr. Qieshi Zhang, Waseda University, Japan
Prof. Hareesha K. S, Manipal University, India
Dr. Emna Fendri, Miracl-FS, Tunisia
Prof. Seokwon Yeom, Daegu University Gyeongsan, Korea
Dr. Elena Battini Sonmez, Istanbul Bilgi University, Turkey
Dr. Szidónia LEFKOVITS, Petru Maior University Tirgu-Mures, Romania
Dr. Ridha Ejbali, National Engineering School of Sfax (ENIS), Tunisia
Dr. Olfa Jemai, University of Gabes, Tunisia
Prof. Mounîm A. El-Yacoubi, University of Paris Saclay, France
Assoc. Prof. Yassin BenAyed, Miracl laboratory, Tunisia

Local Information

Conference Venue

Conference will be held at Fleming´s Conference Hotel Wien,Vienna, Austria. Address: Neubaugürtel 26-28, 1070 Wien, Austria.

Time

UTC/GMT+1

Weather

Vienna has an oceanic climate (hovering just below 22 °C in July and August), and features, according to the Köppen classification, a Cfb (oceanic) -climate. The city has warm summers with average high temperatures of 24 to 33 °C (75 to 91 °F), with maximum exceeding 38 °C (100 °F) and lows of around 17 °C (63 °F). Winters are relatively dry and cold with average temperatures at about freezing point. Spring and autumn are mild. Precipitation is generally moderate throughout the year, averaging 550 mm (21.7 in) annually, with considerable local variations, the Vienna Woods region in the west being the wettest part (700 to 800 mm (28 to 31 in) annually) and the flat plains in the east being the driest part (500 to 550 mm (20 to 22 in) annually). Snow in the winter is not uncommon, but rare compared to Western and Southern regions in Austria.

The Weather Situation of Vienn in November

Average daily minimum temperature	Average daily highest temperature
3.6℃	8.6℃

Currency

Euro

Transportation

Vienna has an extensive transportation network with a unified fare system that integrates municipal, regional and railway systems under the umbrella of the Verkehrsverbund Ost-Region (VOR). Public transport is provided by buses, trams and 5 underground metro lines (U-Bahn), most operated by the Wiener Linien. There are also more than 50 S-trainstations within the city limits. Suburban trains are operated by the ÖBB. The city forms the hub of the Austrian railway system, with services to all parts of the country and abroad. The railway system connects Vienna with other European cities, like Munich, Venice, Budapest, Prague, Bratislava andZürich.

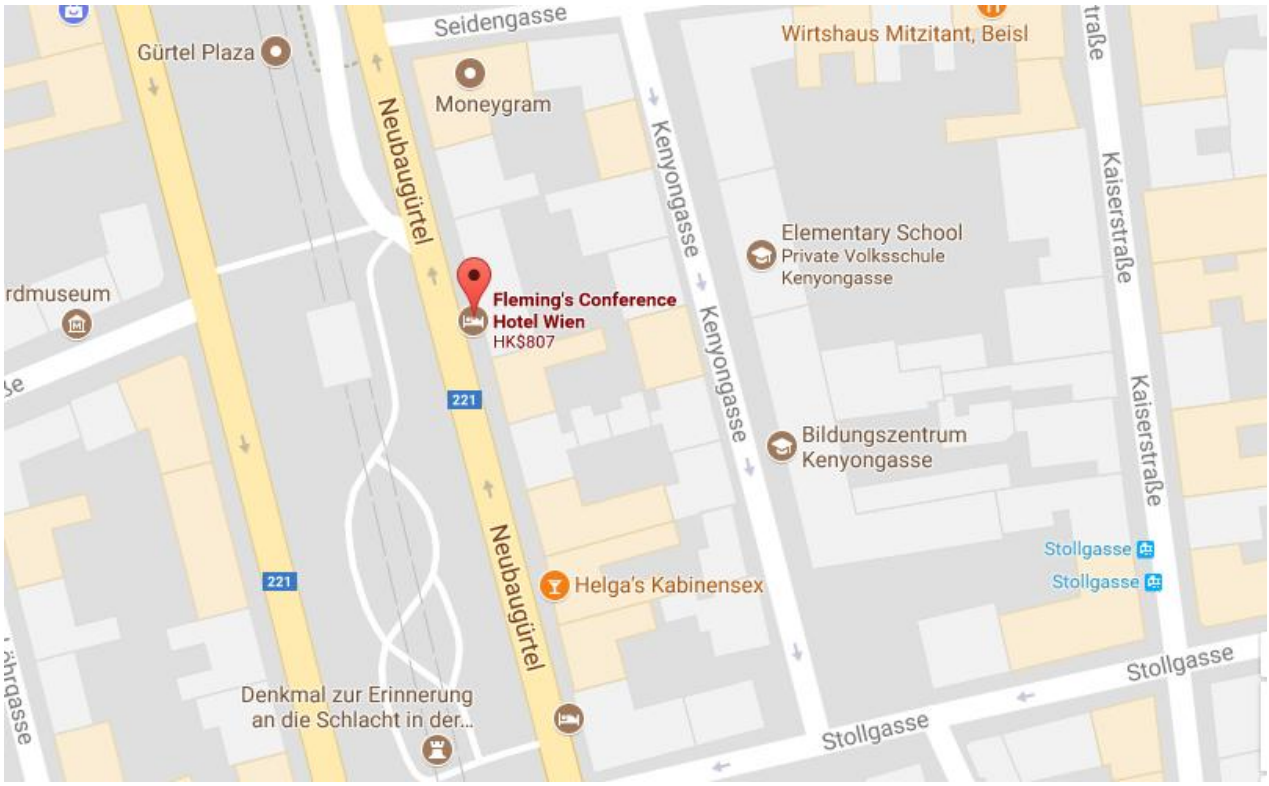
Vienna has multiple road connections including motorways.

Vienna is served by Vienna International Airport, located 18 km (11 mi) southeast of the city centre next to

the town ofSchwechat. The airport handled approximately 22 million passengers in 2014. Following lengthy negotiations with surrounding communities, the airport will be expanded to increase its capacity by adding a third runway. The airport is currently undergoing a major expansion, including a new terminal building opened in 2012 to prepare for an expected increase in passengers.

Access

Fleming's Conference Hotel Wien is only a 5-minute walk from the Westbahnhof Train and Metro station and Vienna's biggest shopping street Mariahilferstraße. The Vienna Airport Bus, leaving every 30 minutes, stops at the Westbahnhof Train Station, and St. Stephen's Cathedral is reachable by a 10-minute metro ride. It will takes about 30 minuts if you by taxi from the international airport to the hotel.



Address: Neubaugürtel 26-28, 1070 Wien, Austria

Instructions for Presentations

Oral Presentations

Oral presentations for ICMV have been allocated 15 minutes of effective presentation time, including Q/A time between Session Chair and speakers.

Authors must prepare their oral presentations to be sure to convey their message in clear and sharp manner, including giving outline of the key principles, facts and results. More detailed discussions can continue during the breaks.

In order to ensure a smooth performance during your session, we kindly ask you to consider the following instructions:

Be at the session room 15 minutes before session starts and introduce yourself to the session chairs.

A video projector and a PC will be available in all conference rooms. Speakers suggested not use their own laptop computer, avoiding useless time breaks in between papers.

Bring your presentation on a USB memory stick in MS-PowerPoint or Adobe PDF formats, and upload it in the Session Room computer no later than 10 minutes prior to your session start! You can also bring it earlier, during the coffee/lunch breaks before your presentation. Please upload your presentation in a right place in order to find it easily at the time of presentation.

Please wear formal clothes or national characteristics of clothing for participation.

In order to avoid any compatibility problems, read carefully the instructions below.

PowerPoint Instructions

For MS-PowerPoint presentations, please use the following versions only: PP 97-2003 (*.ppt) or 2007, 2010 to guarantee that it will be opened successfully on the on-site PC

We recommend to the PPT/PPTX format instead of PPS

All videos or animations in the presentation must run automatically!

Pictures/Videos

We cannot provide support for embedded videos in your presentation; please test your presentation with the on-site PC 20 Minutes before your presentation.

In case your video is not inserted in MS-PowerPoint, it is possible to have it in other formats – MPEG 2,4, AVI (codecs: DivX, XviD, h264) or WMV. Suggested bitrate for all mpeg4 based codecs is about 1 Mbps with SD PAL resolution (1024x576pix with square pixels, AR: 16/9).

Fonts

Only fonts that are included in the basic installation of MS-Windows will be available (English version of Windows). Use of other fonts not included in Windows can cause wrong layout/style of your presentation.

Suggested fonts: Arial, Times New Roman.

If you insist on using different fonts, these must be embedded into your presentation by choosing the right option when saving your presentation:

Click on “File”, then “Save As”

Check the “Tools” menu and select “Embed True Type Fonts”

Poster Presentations

Suggested Poster with size of 60cm*80cm (width*height).

Posters are required to be condensed and attractive. The characters should be large enough so that they are visible from 1 meter apart.

During poster session, the author should stand by your poster, explaining and answering doubts or questions.

Keynote Speakers

Prof. Antanas Verikas

Halmstad University, Sweden

About Prof. Antanas Verikas: He was awarded a PhD degree in pattern recognition from Kaunas University of Technology, Lithuania. Currently he is holding a professor position at both Halmstad University Sweden, where he leads the Department of Intelligent Systems, and Kaunas University of Technology, Lithuania. His research interests include image processing, computer vision, pattern recognition, learning systems, classification, fuzzy logic, and visual media technology. He published more than 150 peer reviewed articles in international journals and conference proceedings and served as Program committee member in numerous international conferences. He is a member of the International Pattern Recognition Society, European Neural Network Society, International Association of Science and Technology for Development, and a member of the IEEE.

Title: Data mining-based prediction of athletic performance

Abstract: Bicycling golf are two kinds of sports considered in this work. Golf shot effectiveness, expressed in terms of club head speed or ball carry distance, and a blood lactate concentration level and an oxygen uptake rate are parameters used to assess performance of golfers and bicyclists in this work. Predicting the performance parameters using information extracted from multi-channel surface electromyography (sEMG) signals recorded from relevant body muscles is the main aim of this work. The approach is based on canonical correlation analysis, random forest-based modelling, decision fusion and multiple sets of novel features extracted from a sEMG signal stream. Statistics reflecting muscle coordination patterns and various statistics computed using correlation-based analysis of sEMG dynamics were important predictors of the parameters. The adopted methodology allowed building accurate models for predicting blood lactate concentration level and oxygen uptake rate using a set of simple features.

Surface EMG data enable wearable computing in the field of ambient intelligence and has potential to enhance exercising of a long carry distance shot. We argue that muscle coordination patterns change during fatiguing cycling exercises and these changes provide important information for tracking two important physiological parameters, blood lactate concentration level and oxygen uptake rate. The investigations have also shown that the canonical correlation analysis is a promising tool for studying relations between sEMG and biomechanical data. Better understanding of these relations may lead to guidelines concerning muscle engagement and coordination of thorax, pelvis and arms during a golf swing and will help golf coaches in providing substantiated advices.

Prof. Petia Radeva

University of Barcelona, Spain

Dr. Petia Radeva (PhD 1993, Universitat Autonoma de Barcelona, Spain) is a senior researcher and associate professor at UB. She received her Ph.D. degree from the Universitat Autonoma de Barcelona in 1998. Her present research interests are on development of learning-based approaches (in particular, statistical methods) for computer vision. She has led one EU project and 12 Spanish projects, 14 patents in the field of computer vision and medical imaging applied to health. Dr. Petia Radeva has published more than 80 SCI-indexed articles and more than 150 international publications. She has an H-index of 29 with 3166 citations. Dr. Petia Radeva was awarded the prize of "Antoni Caparros" for best technology transfer project of University of Barcelona (2013) and the prize of "Jaume Casademont" (2012) for her research activities in the field of lifelogging.

Title: Fusing Convolutional Neural Networks. Application to Food Analysis

Abstract: With the arrival of Convolutional Neural Networks, the complex problem of food recognition has experienced an important improvement recently. The best results have been obtained using methods based on very deep Convolutional Neural Networks, which show that the deeper the model, the better the classification accuracy is. However, very deep neural networks may suffer from the overfitting problem. In this paper, we propose a combination of multiple classifiers based on Convolutional models that complement each other and thus, achieve an improvement in performance. Moreover, we show how a multi-task framework can allow to combine and solve multiple problems of food analysis. The evaluation of our approach is done on 2 public datasets: Food-101 as a dataset with a wide variety of fine-grained dishes, and Food-11 as a dataset of high-level food categories, where our approach outperforms the independent Convolutional Neural Networks models.

Prof. Dr. Petra Perner (IAPR Fellow)

Institute of Computer Vision and Applied Computer Sciences,

Leipzig,Germany

Petra Perner is the director of the Institute of Computer Vision and Applied Computer Sciences in Leipzig, Germany. Under her leadership the institute developed to one of the leading institutes in basic and applied research in computer vision, data mining, machine learning, case-based reasoning and image databases.

She received a diploma degree in electrical engineering and a PhD degree in computer science. Her PhD thesis is about "Data Reduction Methods for Industrial Robots with Direct Teach-in-Programing" and her habilitation thesis is about "A Methodology for the Development of Knowledge-Based Image-Interpretation Systems".

Under her guidance highly qualified national and international personnel got trained in the above mentioned high-technology fields through research projects and education programs.

The tutorial on Data Mining (www.data-mining-tutorial.de) is a recognized tutorial among the domain experts and is especially designed for training industrial people on hot topics in data mining. She gave courses on Data Mining and Image Interpretation at different universities in various countries.

For her contributions to image interpretation by developing methods for case-based reasoning, machine learning and data mining, she was awarded as IAPR Fellow in 2006.

More on: http://www.ibai-research.de/vita.php

Title: Health Monitoring by an automated System for Biological Hazardous Material in the Air

Abstract: Human beings are exposed every day to bioaerosols in the various fields of their personal and/or professional daily life. The European Commission has rules protecting employees in the workplace from biological hazards. Airborne fungi, yeast, bacteria, and viruses can be detected and identified by an image acquisition and interpretation system. In this talk we present results on the development of an image interpretation system fordetecting biological hazardous material in the air. We explain the application domain and describe the development issues. The development strategy and the architecture of the system are described. Finally we give recent results.

Prof. Dmitry Nikolaev

Institute for Information Transmission Problems, RAS (IITP RAS),

Russia

Dmitry Nikolaev was born in Moscow, Russia. He obtained his Master degree (in physics) in 2000 and Ph.D. degree (in computational methods of image processing) in 2004 from Moscow State University. Since 2007 he is a head of the Vision Systems Lab. at the Institute for Information Transmission Problems, RAS. His research activities are in the areas of computer vision with primary application to color image understanding. Since 2005 he gives a course “Image processing” in the Moscow Institute of Physics and Technology, take up an appointment of Assoc. Prof in 2008. At 2015 he has more than 120 publications.

Title: Recognition of documents with fixed layout on a mobile device: coarse-to-fine approach

Abstract: Automatic optical recognition of documents is a traditional function of modern document processing systems. In this field “recognition” means complex process which includes image processing, segmentation, classification and linguistic analysis. Though idea of using mobile devices for recognition of paper documents is not new, direct usage of existing software solutions for scanned images recognition yields low recognition precision on images obtained using a mobile device. This is due, first of all, perspective distortions and lower effective resolution in the latter case. Indeed, one of the most widespread approaches to document images recognition is identification and recognition of arbitrary text on images with subsequent text results analysis. Following this, document classification and fields attribution requires recognition of document template text, which is usually much smaller than filling text and almost unable to read on the mobile images. Super-resolution algorithms allows to obtain a single better quality image, but this approach is too resource-consuming in the sense that it implies improvement of image for the whole document, whereas for a large number of document types we could only focus on several fields regions. In this talk an original approach and algorithms for recognition of video frame sequence containing an image of a document is presented, which is applicable for mobile implementation. It is based on the coarse-to-fine methodology, where template matching and fields localization is performed on the image with lowered resolution, followed by lazy processing of parts of the images only corresponding to the fields which are not recognized yet. Video stream is exploited as a source of noise reduction both in coordinates of the fields and OCR classifiers outputs. The algorithm based on the proposed approach is suitable for running on the device itself and can operate even when none of the video frames contains a document image of sufficient quality by itself.

Prof. Alexander Bernstein

Skolkovo Institute of Science and Technology (Skoltech), Moscow, Russia Federation

Professor Alexander Bernstein is a Principal Research Scientist in Skolkovo Institute of Science and Technology (Skoltech), Center for Computational and Data-Intensive Science and Engineering. Prof. Bernstein started his career at the Research Institute of Automatic Equipment in 1969, where he was developing mathematical models and algorithms for computer networks. At the same time, he was engaged in mathematical statistics. For research in this field, he received the Candidate of Sciences Degree in Math from Steklov Mathematical Institute of USSR Academy of Sciences in 1973 and the Doctor of Sciences degree in Math in 1987 from the Department of Computational Mathematics and Cybernetics of Moscow State University In the year of 1991, the Higher Attestation Commission awarded Alexander with the academic rank of Professor in the field of Intelligent Technologies and Systems. In 2002, Alexander joined the Software Engineering Center of the Russian Academy of Sciences to lead the projects in developing data analysis and applied mathematics software. Prior to joining Skoltech, Alexander held the positions of Chief Researcher at the Institute for System Analysis RAS and Lead Scientist in the Data Analysis Research Lab at the Institute for Information Transmission Problems RAS. At the same time, he had part-time full professor positions at National Research University Higher School of Economics and Moscow Institute of Physics and technology.

Title: Reinforcement Learning in Computer vision

Abstract: Nowadays, machine learning has become a core part of computer vision and one of the basic technologies used in solving various computer vision tasks such as feature detection, image segmentation, object recognition and tracking, etc. Computer vision tasks are formulated as learning problems, and machine learning is an essential and ubiquitous tool for automatic extraction of patterns or regularities from the images.

In many applications, various complex systems such as robots include in their composition various visual sensors from which they learn the state of their environment by solving the corresponding tasks of computer vision. The solutions of these tasks are used in the system for making decisions about its future behavior and, hence, should take into account this use. Reinforcement learning is one of modern machine learning technologies in which learning is carried out through interacting with the environment.

In recent years, Reinforcement learning has been used both for solving the applied tasks included processing and analysis of visual information and solving own problems of computer vision such as filtering, extracting image features, localizing objects in scenes, and many others.

The speech will describe shortly the Reinforcement learning technology and its use for solving the computer vision problems.

Plenary Speaker

Prof. Henryk Palus

Silesian University of Technology, Gliwice, Poland

Henryk Palus received the Ph.D. and D.Sc. (habilitation) degrees in automatic control and robotics from Silesian University of Technology in Gliwice, Poland in 1990 and 2007 respectively. He is currently an associate professor in the Institute of Automatic Control at the Silesian University of Technology. His research interests are mainly in the areas of color image acquisition, color representation, digital image processing, and machine vision. Courses taught include: Vision Sensor Systems, Sensors in Robotics, Visual Information Processing and Robot Vision. In the last fifteen years he co-organized International Conferences on Computer Vision and Graphics (ICCVG) and edited their proceedings. He is/was a member of Advisory Board and reviewer of "Machine Graphics and Vision", founding member and Vice-President of the Polish Association of Image Processing (TPO, IAPR member) and member of the Commission of Metrology of Polish Academy of Sciences. Professor Palus is/was also a member of scientific committees of several conferences in Poland and abroad. He has co-organized several sessions in international and domestic workshops on image processing. He is an author/coauthor of more than 115 papers in international and Polish journals and conference proceedings.

Title: Toward a perceptual image quality assessment of color quantized images

Abstract: Color image quantization is an important operation in the field of color image processing. In this paper, we consider new perceptual image quality metrics for assessment of quantized images. These types of metrics, e.g. DSCSI, MDSIs, MDSIm and HPSI achieve the highest correlation coefficients with MOS during tests on the six publicly available image databases. Research was limited to images distorted by two types of compression: JPG and JPG2K. Statistical analysis of correlation coefficients based on the Friedman test and post-hoc procedures showed that the differences between the four new perceptual metrics are not statistically significant.

Daily Schedule of Events

Nov. 13| Monday	***Lobby***			
	Registration: Amy Hu, Cindy Lau, Jennifer Law ADD: Lobby of Fleming´s Conference Hotel Wien Note: *Collecting conference materials *Accommodation not provided, and it's suggested to make an early reservation.			10:00am-17:00pm
Nov. 13| Monday| afternoon	***Fleming's ROOM 9+10***			
	Opening Remarks: Prof. Petia Radeva, University of Barcelona, Spain			13:00pm-13:05pm
	Keynote Speech I: Prof. Alexander Bernstein, Skolkovo Institute of Science and Technology (Skoltech), Moscow, Russia Federation "Reinforcement learning in computer vision"			13:05pm -13:45pm
	Keynote Speech II: Prof. Petia Radeva, University of Barcelona, Spain " Fusing convolutional neural networks. application to food analysis"			13:45pm -14:25pm
	Coffee Break & Group Photo			14:25pm-14:55pm
	Fleming's ROOM 9	***Fleming's ROOM 10***		
	Session 1-Pattern recognition	Session 2-Target detection and tracking		14:55pm-17:10 pm
Nov. 14| Tuesday| Morning	Opening Remarks: Prof. Antanas Verikas, Halmstad University, Sweden			9:00am-9:05am
	Keynote Speech I: Prof. Dmitry Nikolaev, Institute for Information Transmission Problems, RAS (IITP RAS), Russia "Recognition of documents with fixed layout on a mobile device: coarse-to-fine approach"			9:05am-9:45am
	Keynote Speech II: Prof. Antanas Verikas, Halmstad University, Sweden "Data mining-based prediction of athletic performance"			9:45am-10:25am
	Coffee Break & Group Photo			10:25am-11:00am
	Keynote Speech III: Prof. Dr. Petra Perner (IAPR Fellow),Institute of Computer Vision and Applied Computer Sciences, Leipzig, Germany "Health monitoring by an automated system for biological hazardous material in the air"			11:00am -11:40am
	Lunch @ Fleming's restaurant			12:00am-13:00pm
Nov. 14 | Tuesday | afternoon	***Fleming's ROOM 9***	***Fleming's ROOM 10***	***Fleming's ROOM 6***	
	Session 3 Image transformation and segmentation	Session 4 Video processing and technology	Session 5 Machine vision and visualization	13:00pm-15:30pm

	Coffee Break			15:30pm-15:45pm
	Session 6 Image analysis and imaging system	Session 7 Image processing techniques and methods	Session 8 Information system and image application	15:45pm-18:30pm
	Dinner @ Fleming's restaurant			19:00pm-20:00pm
	Poster Session (Fleming's ROOM 9+10)			13:00pm-18:00pm
Nov. 15	*One day visit*			9:30am-10:30am

Technical Program

Afternoon, Nov.13

Opening & Keynote Speeches

Chair: Prof. Petia Radeva, University of Barcelona, Spain

Venue: Fleming's ROOM 9+10

13:00pm-17: 25pm

13:00pm-13:05pm Opening Remarks	***Fleming's ROOM 9+10***	**Prof. Petia Radeva** University of Barcelona, Spain
13:05pm -13:45pm Keynote Speech I		**Prof. Alexander Bernstein** Skolkovo Institute of Science and Technology (Skoltech), Moscow, Russia Federation "Reinforcement Learning in Computer vision"
13:45pm -14:25 pm Keynote Speech II		**Prof. Petia Radeva** University of Barcelona, Spain "Fusing onvolutional Neural Networks. Application to Food Analysis"
14:25pm-14:55pm		Coffee Break & Group Photo

*The Group Photo will be updated online.

**One best presentation will be selected from each session, the best one will be announced at the end of each session and awarded certificate during the dinner, the winners' photos will be updated online.

***Best Presentation will be evaluated from: Originality; Applicability; Technical Merit; PPT; English.

****Please arrive at the conference room 10 minutes earlier before the session starts, copy your PPT to the laptop.

Morning, Nov.14

Opening & Keynote Speeches

Chair: Prof. Antanas Verikas, Halmstad University, Sweden

Venue: Fleming's ROOM 9+10

9:00am-11:40am

9:00am-9:05am Opening Remarks	***Fleming's ROOM 9+10***	**Prof. Antanas Verikas** Halmstad University, Sweden
9:05am-9:45am Keynote Speech I		**Prof. Dmitry Nikolaev** Institute for Information Transmission Problems, RAS (IITP RAS), Russia "Recognition of documents with fixed layout on a mobile device: coarse-to-fine approach"
9:45am-10:25am Keynote Speech II		**Prof. Antanas Verikas** Halmstad University, Sweden "Data mining-based prediction of athletic performance"
10:25am-11:00am		Coffee Break & Group Photo
11:00am-11:40am		**Prof. Dr. Petra Perner (IAPR Fellow)** Institute of Computer Vision and Applied Computer Sciences, Leipzig, Germany "Health Monitoring by an automated System for Biological Hazardous Material in the Air"
12:00am-13:00pm		Lunch @ Fleming's restaurant

*The Group Photo will be updated online.

**One best presentation will be selected from each session, the best one will be announced at the end of each session and awarded certificate during the dinner, the winners' photos will be updated online.

***Best Presentation will be evaluated from: Originality; Applicability; Technical Merit; PPT; English.

****Please arrive at the conference room 10 minutes earlier before the session starts, copy your PPT to the laptop.

Oral Presentations

Afternoon, Nov. 13

Session 1< Pattern recognition>

Venue: Fleming's ROOM 9
Chair: Prof. Alexander Bernstein, Skolkovo Institute of Science and Technology (Skoltech), Moscow, Russia Federation
Time: 14:55pm-17:10 pm

Note:
* The certificate of oral presentation will be awarded at the end of each session.
* For the Best Presentation of each session, it is encouraged to award it to student author prior. The best presentation will be announced at the end of session and awarded certificate over the dinner banquet.
*Please arrive the meeting room at least 10 Mins before session starts. To show the respect to other authors, especially to encourage the student authors, we strongly suggest you attend the whole session, the scheduled time for presentations might be changed due to unexpected situations, please come as early as you could.
* Session photos will be taken at the end of the session and update online

V0006

Time: 14:55-15:10

Analysis of straw row in the image to control the trajectory of the agricultural combine harvester

Aleksandr Yurievich Shkanaev, Dmitry Valerevich Polevoy, Aleksei Vladimirovich Panchenko, **Ms. Krokhina Darya Alekseevna**, Sadekov Rinat Nailevish
Cognitive Technologies LTD, Russian Federation

The paper proposes a solution to the automatic operation of the combine harvester along the straw rows by means of the images from the camera, installed in the cab of the harvester. The U-Net is used to recognize straw rows in the image. The edges of the row are approximated in the segmented image by the curved lines and further converted into the harvester coordinate system for the automatic operating system. The "new" network architecture and approaches to the row approximation has improved the quality of the recognition task and the processing speed of the frames up to 96% and 7.5 fps, respectively.

V0029

Time: 15:10-15:25

Comparision of Lpc Based Parametric Techniques For Respiratory Sounds Recognition

Ms. Fatma Zehra Göğüş and Gülay Tezel
Selcuk University, Turkey

Respiratory sounds are widely adopted marker of several diseases associated with upper and lower respiratory systems and lungs. Hence, recognition of respiratory sounds is an important step in diagnosis of the several diseases. In this study, it is aimed to recognize normal and asthmatic respiratory sounds. To accomplish this aim, analysis and classification process of the sounds were performed. LPC-based parametric techniques namely Linear Predictive Coefficients (LPC), Linear Prediction Cepstral Coefficients (LPCC) and Weighted Linear Prediction Cepstral Coefficients (WLPCC) techniques were used in analysis and feature extraction process. Linear prediction coefficients, cepstral coefficients and weighted cepstral

coefficients were evaluated as characteristic features of the sound signals. In addition, Fuzzy C Means (FCM) clustering algorithm was used to achieve feature reduction. k nearest neighbor (kNN) and fuzzy k nearest neighbor (F-kNN) classifiers were design to classify respiratory sounds as normal and asthmatic sound signals. As a result of this study, the LPC-based parametric techniques were compared in terms of the effect on classification.

V0035

Time: 15:25-15:40

Geometry-Based Populated Chessboard Recognition

Youye Xie, Gongguo Tang and **Dr. William Hoff**
DAQRI Holographics Ltd, Austria

Chessboards are commonly used to calibrate cameras, and many robust methods have been developed to recognize the unoccupied boards. However, when the chessboard is populated with chess pieces, such as during an actual game, the problem of recognizing the board is much harder. Challenges include occlusion caused by the chess pieces, the presence of outlier lines and low viewing angles of the chessboard. In this paper, we present a novel approach to address the above challenges and recognize the chessboard. The Canny edge detector and Hough transform are used to capture all possible lines in the scene. The k-means clustering and a k-nearest-neighbors inspired algorithm are applied to cluster and reject the outlier lines based on their Euclidean distances to the nearest neighbors in a scaled Hough transform space. Finally, based on prior knowledge of the chessboard structure, a geometric constraint is used to find the correspondences between image lines and the lines on the chessboard through the homography transformation. The proposed algorithm works for a wide range of the operating angles and achieves high accuracy in experiments.

V0046

Time: 15:40-15:55

Deep Learning based Beat Event Detection in Action Movie Franchises

Dr. Naveed Ejaz, Umair Ali Khan, Miguel Ángel Mart ńez-Del-Amor,Heiko Sparenberg
Fraunhofer Institute of Integrated Circuits, Erlangen, Germany

Automatic understanding and interpretation of movies can be used in a variety of ways to semantically manage the massive volumes of movies data. “Action Movie Franchises” dataset is a collection of twenty Hollywood action movies from five famous franchises with ground truth annotations at shot and beat level of each movie. In this dataset, the annotations are provided for eleven semantic beat categories. In this work, we propose a deep learning based method to classify shots and beat-events on this dataset. The training dataset for each of the eleven beat categories is developed and then a Convolution Neural Network is trained. After finding the shot boundaries, key frames are extracted for each shot and then three classification labels are assigned to each key frame. The classification labels for each of the key frames in a particular shot are then used to assign a unique label to each shot. A simple sliding window based method is then used to group adjacent shots having the same label in order to find a particular beat event. The results of beat event classification are presented based on criteria of precision, recall, and F-measure. The results are compared with the existing technique and significant improvements are recorded.

V0050

Time: 15:55-16:10

Local Feature Extraction from RGB and Depth Videos for Human Action Recognition

Ms. Rawya Al-Akam, Prof. Dr. Ing. Deitrich Paulus
University of Koblenz.Landau, Germany

In this paper, we present a novel system to analyze human body motions for action recognition tasks where we use 3D videos (RGB and depth data). We apply the Bag-of-Features approach for recognizing human actions by extracting local-spatial temporal features from all video frames. Feature vectors are computed in two steps: The first step consists of detecting all interest keypoints from RGB video frames using Speed-Up Robust Features; we filters motion points using Motion History Image and Optical Flow, then aligned these motion points to the depth frame sequences. The second step uses a Histogram of Orientation Gradient descriptor for computing the features vector around these points from both RGB and depth channels, then combines these feature values in one RGBD feature vector. Finally, we used k-means clustering and multi-class Support Vector Machines for the classification task. Our system is invariant to scale, rotation and illumination. All tested results are computed from a dataset that is available to the public and often used in the community. By using our new feature combination method, we reach recognition rates superior to other publications on the dataset.

V0052

Time: 16:10-16:25

Speaker Emotion Recognition: From Classical Classifiers To Deep Neural Networks

Eya Mezghani, Maha Charfeddine, **Prof. Henri Nicolas**, and Chokri Ben Amar
University of Bordeaux, France

Speaker emotion recognition is considered among the most challenging tasks in recent years. In fact, automatic systems for security, medicine or education can be improved when considering the speech affective state. In this paper, a twofold approach for speech emotion classification is proposed. At the first side, a relevant set of features is adopted, and then at the second one, numerous supervised training techniques, involving classic methods as well as deep learning, are experimented. Experimental results indicate that deep architecture can improve classification performance on two affective databases, the Berlin Dataset of Emotional Speech and the SAVEE Dataset Surrey Audio-Visual Expressed Emotion.

V0069

Time: 16:25-16:40

Biometric Identification Based on Feature Fusion with PCA

Assoc. Prof. László Lefkovits, Szidónia Lefkovits and Simina Emerich
Sapientia University Tg. Mures, Romania

Biometric identification is gaining ground compared to traditional identification methods. Many biometric measurements may be used for secure human identification. The most reliable among them is the iris pattern because of its uniqueness, stability, unforgeability and inalterability over time. The approach presented in this paper is a fusion of different feature descriptor methods such as HOG, LIOP, LBP, used for extracting iris texture information. The classifiers obtained through the SVM and PCA methods demonstrate the effectiveness of our system applied to one and both irises. The performances measured are highly accurate and foreshadow a fusion system with a rate of identification approaching 100% on the UPOL database.

V0093

Time: 16:40-16:55

Comparison of the scanned pages of the contractual documents

Elena Andreeva, Vladimir V. Arlazarov, **Temudzhin Manzhikov**, Oleg Slavin
Moscow Institute of Physics and Technology, Moscow, Russia

In this paper the problem statement is given to compare the digitized pages of the official papers. Such problem appears during the comparison of two customer copies signed at different times between two parties with a view to find the possible modifications introduced on the one hand. This problem is a practically significant in the banking sector during the conclusion of contracts in a paper format. The method of comparison based on the recognition, which consists in the comparison of two bag-of-words, which are the recognition result of the master and test pages, is suggested. The described experiments were conducted using the OCR Tesseract and the siamese neural network. The advantages of the suggested method are the steady operation of the comparison algorithm and the high exacting precision, and one of the disadvantages is the dependence on the chosen OCR.

V0112

Time: 16:55-17:10

Textual Blocks Rectification Method Based on Fast Hough Transform Analysis in Identity Documents Recognition

Mr. Bezmaternykh P.V., Nikolaev D.P., Arlazarov V.L.
Smart Engines, Russia

Textual blocks rectification or slant correction is an important stage of document image processing in OCR systems. This paper considers existing methods and introduces an approach for the construction of such algorithms based on Fast Hough Transform analysis. A quality measurement technique is proposed and obtained results are shown for both printed and handwritten textual blocks processing as a part of an industrial system of identity documents recognition on mobile devices.

Session 2< Target detection and tracking>

Venue: Fleming's ROOM 10
Chair: Prof. Petia Radeva, University of Barcelona, Spain
Time: 14:50pm-17:10pm

Note:
* The certificate of oral presentation will be awarded at the end of each session.
* For the Best Presentation of each session, it is encouraged to award it to student author prior. The best presentation will be announced at the end of session and awarded certificate over the dinner banquet.
*Please arrive the meeting room at least 10 Mins before session starts. To show the respect to other authors, especially to encourage the student authors, we strongly suggest you attend the whole session, the scheduled time for presentations might be changed due to unexpected situations, please come as early as you could.
* Session photos will be taken at the end of the session and update online

V0013

Time: 14:55-15:10

Automated Grain Extraction and Classification by Combining Improved Region Growing Segmentation and Shape Descriptors in Electromagnetic Mill Classification System

Dr. Sebastian Budzan
Silesian University of Technology, Gliwice, Poland

In this paper, the automatic method of grain detection and classification has been presented. As input, it uses a single digital image obtained from milling process of the copper ore with an high-quality digital camera. The grinding process is an extremely energy and cost consuming process, thus granularity evaluation process should be performed with high efficiency and time consumption. The method proposed in this paper is based on the three-stage image processing. First, using Seeded Region Growing (SRG) segmentation with proposed adaptive thresholding based on the calculation of Relative Standard Deviation (RSD) all grains are detected. In the next step results of the detection are improved using information about the shape of the detected grains using distance map. Finally, each grain in the sample is classified into one of the predefined granularity class. The quality of the proposed method has been obtained by using nominal granularity samples, also with a comparison to the other methods.

V0057

Time: 15:10-15:25

Aggregated Channels Network for Real-Time Pedestrian Detection

Mr. Farzin Ghorban, Javier Mar ń, Yu Su, Alessandro Colombo, Anton Kummert
University of Wuppertal, Germany

Convolutional neural networks (CNNs) have demonstrated their superiority in numerous computer vision tasks, yet their computational cost results prohibitive for many real-time applications such as pedestrian detection which is usually performed on low-consumption hardware. In order to alleviate this drawback, most strategies focus on using a two-stage cascade approach. Essentially, in the first stage a fast method generates a significant but reduced amount of high quality proposals that later, in the second stage, are evaluated by the CNN. In this work, we propose a novel detection pipeline that further benefits from the two-stage cascade strategy. More concretely, the enriched and subsequently compressed features used in the first stage are reused as the CNN input. As a consequence, a simpler network architecture, adapted for such small input sizes, allows to achieve real-time performance and obtain results close to the state-of-the-art while running significantly faster without the use of GPU. In particular, considering that the proposed pipeline runs in frame rate, the achieved performance is

highly competitive. We furthermore demonstrate that the proposed pipeline on itself can serve as an effective proposal generator.

V0070

Time: 15:25-15:40

Early detection of lung cancer from CT images: Nodule segmentation and classification using deep learning

Manu Sharma, **Assist. Prof. Jignesh S. Bhatt**, Manjunath V. Joshi
Indian Institute of Information Technology, Vadodara (INDIA)

Lung cancer is one of the most abundant causes of the cancerous deaths worldwide. It has low survival rate mainly due to the late diagnosis. With the hardware advancements in computed tomography (CT) technology, it is now possible to capture the high resolution images of lung region. However, it needs to be augmented by efficient algorithms to detect the lung cancer in the earlier stages using the acquired CT images. To this end, we propose a two-step algorithm for early detection of lung cancer. Given the CT image, we first extract the patch from the center location of the nodule and segment the lung nodule region. We propose to use Otsu method followed by morphological operations for the segmentation. This step enables accurate segmentation due to the use of data-driven threshold. Unlike other methods, we perform the segmentation without using the complete contour information of the nodule. In the second step, a deep convolutional neural network (CNN) is used for the better classification (malignant or benign) of the nodule present in the segmented patch. Accurate segmentation of even a tiny nodule followed by better classification using deep CNN enables the early detection of lung cancer. Experiments have been conducted using 6306 CT images of LIDC-IDRI database. We achieved the test accuracy of 84.13%, with the sensitivity and specificity of 91.69% and 73.16%, respectively, clearly outperforming the state-of-the-art algorithms.

V0108

Time: 15:40-15:55

Vision Based Speed Breaker Detection for Autonomous Vehicle

Arvind C S, **Mr. Ritesh Mishra**, Kumar Vishal, Venugopal Gundimeda
Cognizant Technology Solutions Pvt. Ltd, India

In this paper, we are presenting a robust and real-time, vision-based approach to detect speed breaker in urban environments for autonomous vehicle. Our method is designed to detect the speed breaker using visual inputs obtained from a camera mounted on top of a vehicle. The method performs inverse perspective mapping to generate top view of the road and segment out region of interest based on difference of Gaussian and median filter images. Furthermore, the algorithm performs RANSAC line fitting to identify the possible speed breaker candidate region. This initial guessed region via RANSAC, is validated using support vector machine. Our algorithm can detect different categories of speed breakers on cement, asphalt and interlock roads at various conditions and have achieved a recall of ~0.98.

V0130

Time: 15:55-16:10

Multi person detection and tracking based on hierarchical Level-Set Method

Dr. Chadia Khraief, Faouzi Benzarti and Hamid Amiri
University of Tunis El Manar, National Engineering School of Tunis (ENIT), Signal, Image and Technology Information (LR-SITI), Manar, Tunisia

In this paper, we propose an efficient unsupervised method for mutli-person tracking based on hierarchical level-set approach. The proposed method uses both edge and region information in order to effectively detect objects. The persons are tracked on each frame of the sequence by minimizing an energy functional that combines color, texture and shape information. These features are enrolled in covariance matrix as region descriptor. The present method is fully automated without the need to manually specify the initial contour of Level-set. It is based on combined person detection and background subtraction methods. The edge-based is employed to maintain a stable evolution, guide the segmentation towards apparent boundaries and inhibit regions fusion. The computational cost of level-set is reduced by using narrow band technique. Many experimental results are performed on challenging video sequences and show the effectiveness of the proposed method.

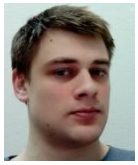
V0138

Time: 16:10-16:25

Vanishing Points Detection Using Combination of Fast Hough Transform and Deep Learning

Mr. Alexander Sheshkus, Anastasia Ingacheva and Dmitry Nikolaev
Smart Engines, Russia

In this paper we propose a novel method for vanishing points detection based on convolutional neural network (CNN) approach and Fast Hough transform algorithm. We show how to determine Fast Hough Transform neural network layer and how to use it in order to increase usability of the neural network approach to the vanishing point detection task. Our algorithm includes CNN with consequence of convolutional and fast Hough transform layers. We are building estimator for distribution of possible vanishing points in the image. This distribution can be used to find candidates of vanishing point. We provide experimental results from tests of suggested method using images collected from videos of a road trips. Our approach shows stable result on test images with different projective distortions and noise. Described approach can be effectively implemented for mobile GPU and CPU.



V0116

Time: 16:25-16:40

Algorithm for Covert Convoy of a Moving Target Using a Group of Autonomous Robots

Igor Polyakov, **Dr. Evgeny Shvets**
Institute for Information Transmission Problems RAS

An important application of autonomous robot systems is to substitute human personnel in dangerous environments to reduce their involvement and subsequent risk on human lives. In this paper we solve the problem of covertly convoying a civilian in a dangerous area with a group of unmanned ground vehicles (UGVs) using social potential fields. The novelty of our work lies in the usage of UGVs as compared to the unmanned aerial vehicles typically employed for this task in the approaches described in literature. Additionally, in our paper we assume that the group of UGVs should simultaneously solve the problem of patrolling to detect intruders on the area. We develop a simulation system to test our algorithms, provide numerical results and give recommendations on how to tune the potentials governing robots' behaviour to prioritize between patrolling and convoying tasks.

V0109

Time: 16:40-16:55

Deep Learning Architecture for Recognition of Abnormal Activities

Marwa Khatrouch, **Ms. Mariem Gnouma**, Ridha Ejbali and Mourad Zaied

Gabes Tunisia

The video surveillance is one of the key areas in computer vision researches. The scientific challenge in this field involves the implementation of automatic systems to obtain detailed information about individuals and groups behaviors. In particular, the detection of abnormal movements of groups or individuals requires a fine analysis of frames in the video stream. In this article, we propose a new method to detect anomalies in crowded scenes. We try to categorize the video in a supervised mode accompanied by unsupervised learning using the principle of the autoencoder. In order to construct an informative concept for the recognition of these behaviors, we use a technique of representation based on the superposition of human silhouettes. The evaluation of the UMN dataset demonstrates the effectiveness of the proposed approach.

V0002

Time: 16:55-17:10

Gender classification from face images by using Local Binary Pattern and Gray-Level Co-Occurrence Matrix

Ms. Betül UZBAŞ and Ahmet ARSLAN
Sel çuk University, Konya, Turkey

Gender is an important step for human computer interactive processes and identification. Human face image is one of the important sources to determine gender. In the present study, gender classification is performed automatically from facial images.
In order to classify gender, we propose a combination of features that have been extracted face, eye and lip regions by using a hybrid method of Local Binary Pattern and Gray-Level Co-Occurrence Matrix. The features have been extracted from automatically obtained face, eye and lip regions. All of the extracted features have been combined and given as input parameters to classification methods (Support Vector Machine, Artificial Neural Networks, Naive Bayes and k-Nearest Neighbor methods) for gender classification. The Nottingham Scan face database that consists of the frontal face images of 100 people (50 male and 50 female) is used for this purpose. As the result of the experimental studies, the highest success rate has been achieved as 98% by using Support Vector Machine. The experimental results illustrate the efficacy of our proposed method.

Afternoon, Nov. 14

Session 3<Image transformation and segmentation >

Venue: Fleming's ROOM 9
Chair: Prof. Antanas Verikas, Halmstad University, Sweden
Time: 13:00pm-15:15pm

Note:
* The certificate of oral presentation will be awarded at the end of each session.
* For the Best Presentation of each session, it is encouraged to award it to student author prior. The best presentation will be announced at the end of session and awarded certificate over the dinner banquet.
*Please arrive the meeting room at least 10 Mins before session starts. To show the respect to other authors, especially to encourage the student authors, we strongly suggest you attend the whole session, the scheduled time for presentations might be changed due to unexpected situations, please come as early as you could.
* Session photos will be taken at the end of the session and update online

V0060

Time: 13:00-13:15

Video Segmentation using Keywords

Mr. Vinh Ton-That, Chi-Tai Vong, Xuan-Truong Nguyen-Dao, Minh-Triet Tran
University of Science, VNU-HCM, Vietnam

At DAVIS-2016 Challenge, many state-of-art video segmentation methods achieve potential results, but they still much depend on annotated frames to distinguish between background and foreground. It takes a lot of time and efforts to create these frames exactly. In this paper, we introduce a method to segment objects from video based on keywords given by user. First, we use a real-time object detection system - YOLOv2 to identify regions containing objects that have labels match with the given keywords in the first frame. Then, for each region identified from the previous step, we use Pyramid Scene Parsing Network to assign each pixel as foreground or background. These frames can be used as input frames for Object Flow algorithm to perform segmentation on entire video. We conduct experiments on a subset of DAVIS-2016 dataset in half the size of its original size, which shows that our method can handle many popular classes in PASCAL VOC 2012 dataset with acceptable accuracy, about 75.03%. We suggest widely testing by combining other methods to improve this result in the future.

V0077

Time: 13:15-13:30

A Novel Automatic Segmentation Workflow of Axial Breast DCE-MRI

Ms. Feten Besbes, Norhene Gargouri, Alima Damak, Dorra Sellami
Computer, Imaging, Electronics & Systems (CIELS) from CEM-LAB
National Engineering School of Sfax, Tunisia

In this paper we propose a novel process of a fully automatic breast tissue segmentation which is independent from expert calibration and contrast. The proposed algorithm is composed by two major steps. The first step consists in the detection of breast boundaries. It is based on image content analysis and Moore-Neighbour tracing algorithm. As a processing step, Otsu thresholding and neighbors algorithm are applied. Then, the external area of breast is removed to get an approximated breast region. The second preprocessing step is the delineation of the chest wall which is considered as the lowest cost path linking three key points; These points are located automatically at the breast. They are respectively, the left and right boundary points and the

middle upper point placed at the sternum region using statistical method. For the minimum cost path search problem, we resolve it through Dijkstra algorithm. Evaluation results reveal the robustness of our process face to different breast densities, complex forms and challenging cases. In fact, the mean overlap between manual segmentation and automatic segmentation through our method is 96.5%. A comparative study shows that our proposed process is competitive and faster than existing methods. The segmentation of 120 slices with our method is achieved at least in 20.57±5.2s.

V0084

Time: 13:30-13:45

Breast mass segmentation in mammograms combining Fuzzy C-Means and active contours

Ms. Marwa Hmida, Kamel Hamrouni, Basel Solaiman, Sana Boussetta
Université de Tunis El Manar, Ecole Nationale d'Ingénieurs de Tunis, LR−SITI Signal Image et Technologie de l'information, Tunis, Tunisie

Segmentation of breast masses in mammograms is a challenging issue due to the nature of mammography and the characteristics of masses. In fact, mammographic images are poor in contrast and breast masses have various shapes and densities with fuzzy and ill-defined borders. In this paper, we propose a method based on a modified Chan-Vese active contour model for mass segmentation in mammograms. We conduct the experiment on mass Regions of Interest (ROI) extracted from the MIAS database. The proposed method consists of mainly three stages: Firstly, the ROI is preprocessed to enhance the contrast. Next, two fuzzy membership maps are generated from the preprocessed ROI based on fuzzy C-Means algorithm. These fuzzy membership maps are finally used to modify the energy of the Chan-Vese model and to perform the final segmentation. Experimental results indicate that the proposed method yields good mass segmentation results.

V0088

Time: 13:45-14:00

Squeeze-SegNet: A new fast Deep Convolutional Neural Network for Semantic Segmentation

Mr. Geraldin NANFACK, Azeddine ELHASSOUNY, Rachid OULAD HAJ THAMI
RABAT IT CENTER, ENSIAS, University of Mohamed 5 in Rabat, Morocco

The recent researches in Deep Convolutional Neural Network have focused their attention on improving accuracy that provide significant advances. However, if they were limited to classification tasks, nowadays with contributions from Scientific Communities who are embarking in this field, they have become very useful in higher level tasks such as object detection and pixel-wise semantic segmentation. Thus, brilliant ideas in the field of semantic segmentation with deep learning have completed the state of the art of accuracy, however this architectures become very difficult to apply in embedded systems as is the case for autonomous driving. We present a new Deep fully Convolutional Neural Network for pixel-wise semantic segmentation which we call Squeeze-SegNet. The architecture is based on Encoder-Decoder style. We use a SqueezeNet-like encoder and a decoder formed by our proposed squeeze-decoder module and upsample layer using downsample indices like in SegNet and we add a deconvolution layer to provide final multi-channel feature map. On datasets like Camvid or City-states, our net gets SegNet-level accuracy with less than 10 times fewer parameters than SegNet.

V0121

Time: 14:00-14:15

Modification of YAPE keypoint detection algorithm for wide local contrast range images

Mr. A. Lukoyanov, D. Nikolaev, and I. Konovalenko
Institute for Information Transmission Problems (IITP RAS)

Keypoint detection is an important tool of image analysis, and among many contemporary keypoint detection algorithms YAPE is known for its computational performance, allowing its use in mobile and embedded systems. One of its shortcomings is high sensitivity to local contrast which leads to high detection density in high-contrast areas while missing detections in low-contrast ones.
In this work we study the contrast sensitivity of YAPE and propose a modification which compensates for this property on images with wide local contrast range (Yet Another Contrast-Invariant Point Extractor, YACIPE). As a model example, we considered the traffic sign recognition problem, where some signs are well-lighted, whereas others are in shadows and thus have low contrast. We show that the number of traffic signs on the image of which has not been detected any keypoints is 40% less for the proposed modification compared to the original algorithm.

V0044

Time: 14:15-14:30

Establishing the Correspondence Between Closed Contours of Objects

Alexey V. Savchik, **Assoc. Prof. Victoria A. Sablina** and Dmitry P. Nikolaev
Ryazan State Radio Engineering University (RSREU), Russia

In this paper, we consider the task of finding the correspondence between closed contours of objects in an image pair with small projective distortions. Several methods are considered and their comparison is performed. The experiments results for two contour sets are provided. Sufficient conditions of the applicability of the method of selecting the nearest contour are represented and proven.

V0054

Time: 14:30-14:45

3-D shape recovery from Image focus using Gabor Features

Mr. Fahad Mahmood, Jawad Mahmood and Javaid Iqbal
NUST, Pakistan

Recovering an accurate and precise depth map from a set of acquired 2-D image dataset of the target object each having different focus information is an ultimate goal of 3-D shape recovery. Focus measure algorithm plays an important role in this architecture as it converts the corresponding color value information into focus information which will be then utilized for recovering depth map. This article introduces Gabor features as focus measure approach for recovering depth map from a set of 2-D images. Frequency and orientation representation of Gabor filter features is similar to human visual system and normally applied for texture representation. Due to its little computational complexity, sharp focus measure curve, robust to random noise sources and accuracy, it is considered as superior alternative to most of recently proposed 3-D shape recovery approaches. This algorithm is deeply investigated on real image sequences and synthetic image dataset. The efficiency of the proposed scheme is also compared with the state of art 3-D shape recovery approaches. Finally, by means of two global statistical

measures, root mean square error and correlation, we claim that this approach –in spite of simplicity- generates accurate results.

V0079

Time: 14:45-15:00

Neural Network-based Feature Point Descriptors for Registration of Optical and SAR Images

Dmitry Abulkhanov, Ivan Konovalenko, **Mr. Dmitry Nikolaev Alexey Savchik**, Evgeny Shvets, Dmitry Sidorchuk
IITP RAS, 127051, Moscow, Russia

Registration of images of di_erent nature is an important technique used in image fusion, change detection, e_cient information representation and other problems of computer vision. Solving this task using featurebased approaches is usually more complex than registration of several optical images because traditional feature descriptors (SIFT, SURF, etc.) perform poorly when images have di_erent nature. In this paper we consider the problem of registration of SAR and optical images. We train neural network to build feature point descriptors and use RANSAC algorithm to align found matches. Experimental results are presented that con_rm the method's e_ectiveness.

V0098

Time: 15:00-15:15

Mobile and Embedded Fast High Resolution Image Stitching for Long Length Rectangular Monochromatic Objects with Periodic Structure

Ms. Elena Limonova, Daniil Tropin, Boris Savelyev, Igor Mamay, Dmitry Nikolaev
Institute for Systems Analysis, FRC CSC RAS, Russia

In this paper we describe stitching protocol, which allows to obtain high resolution images of long length monochromatic objects with periodic structure. This protocol can be used for long length documents or human-induced objects in satellite images of uninhabitable regions like Arctic regions. The length of such objects can reach notable values, while modern camera sensors have limited resolution and are not able to provide good enough image of the whole object for further processing, e.g. using in OCR system. The idea of the proposed method is to acquire a video stream containing full object in high resolution and use image stitching. We expect the scanned object to have straight boundaries and periodic structure, which allow us to introduce regularization to the stitching problem and adapt algorithm for limited computational power of mobile and embedded CPUs. With the help of detected boundaries and structure we estimate homography between frames and use this information to reduce complexity of stitching. We demonstrate our algorithm on mobile device and show image processing speed of 2 fps on Samsung Exynos 5422 processor.

Session 4< Video processing and technology>

Venue: Fleming's ROOM 10

Chair: Prof. Dr. Petra Perner (IAPR Fellow), Institute of Computer Vision and Applied Computer Sciences, Leipzig, Germany

Time: 13:00pm-15:30pm

Note:
* The certificate of oral presentation will be awarded at the end of each session.
* For the Best Presentation of each session, it is encouraged to award it to student author prior. The best presentation will be announced at the end of session and awarded certificate over the dinner banquet.
*Please arrive the meeting room at least 10 Mins before session starts. To show the respect to other authors, especially to encourage the student authors, we strongly suggest you attend the whole session, the scheduled time for presentations might be changed due to unexpected situations, please come as early as you could.
* Session photos will be taken at the end of the session and update online

V0092

Time: 13:00-13:15

Method of determining the necessary number of observations for video stream documents recognition

Vladimir V. Arlazarov, Konstantin Bulatov, Oleg Slavin, Igor Janiszewski, **Mr. Temudzhin Manzhikov**
Moscow Institute of Physics and Technology, Moscow, Russia

This paper discusses a task of document recognition on a sequence of video frames. In order to optimize the processing speed an estimation is performed of stability of recognition results obtained from several video frames. Considering identity document (Russian internal passport) recognition on a mobile device it is shown that significant decrease is possible of the number of observations necessary for obtaining precise recognition result.

V0119

Time: 13:15-13:30

Image quality assessment for video stream recognition systems

Mr. Timofey Chernov, Nikita Razumnuy, Alexander Kozharinov, Dmitry Nikolaev, Vladimir Arlazarov,
Federal Research Centre "Computer Science and Control Systems" The Institute for Systems Analysis of Russian Academy of Sciences, Moscow, Russia

Recognition and machine vision systems have long been widely used in many disciplines to automate various processes of life and industry. Input images of optical recognition systems can be subjected to a large number of different distortions, especially in uncontrolled or natural shooting conditions, which leads to unpredictable results of recognition systems, making it impossible to assess their reliability. For this reason, it is necessary to perform quality control of the input data of recognition systems, which is facilitated by modern progress in the field of image quality evaluation. In this paper, we investigate the approach to designing optical recognition systems with built-in input image quality estimation modules and feedback, for which the necessary definitions are introduced and a model for describing such systems is constructed. The efficiency of this approach is illustrated by the example of solving the problem of selecting the best frames for recognition in a video stream for a system with limited resources. Experimental results are presented for the system for identity documents recognition, showing a significant increase in the accuracy and speed of the system under simulated conditions of automatic camera focusing,

leading to blurring of frames.

V0012

Time: 13:30-13:45

Recurrent Neural Network Based Virtual Detection Line

Mr. Roberts Kadikis
EDI (Institute of Electronics and Computer Science), Latvia

The paper proposes an efficient method for detection of moving objects in the video. The objects are detected when they cross a virtual detection line. Only the pixels of the detection line are processed, which makes the method computationally efficient. A Recurrent Neural Network processes these pixels. The machine learning approach allows one to train a model that works in different and changing outdoor conditions. Also, the same network can be trained for various detection tasks, which is demonstrated by the tests on vehicle and people counting. In addition, the paper proposes a method for semi-automatic acquisition of labeled training data. The labeling method is used to create training and testing datasets, which in turn are used to train and evaluate the accuracy and efficiency of the detection method. The method shows similar accuracy as the alternative efficient methods but provides greater adaptability and usability for different tasks.

V0089

Time: 13:45-14:00

Performance Improvement of Multi-class Detection Using Greedy Algorithm for Viola-Jones Cascade Selection

Alexander A. Tereshin, **Mr. Sergey A. Usilin**, Vladimir V. Arlazarov
Smart Engines Ltd., Russia

This paper aims to study the problem of multi-class object detection in video stream with Viola-Jones cascades. An adaptive algorithm for selecting Viola-Jones cascade based on greedy choice strategy in solution of the N-armed bandit problem is proposed. The efficiency of the algorithm on the problem of detection and recognition of the bank card logos in the video stream is shown. The proposed algorithm can be effectively used in documents localization and identification, recognition of road scene elements, localization and tracking of the lengthy objects , and for solving other problems of rigid object detection in a heterogeneous data flows. The computational efficiency of the algorithm makes it possible to use it both on personal computers and on mobile devices based on processors with low power consumption.

V0063

Time: 14:00-14:15

First stereo video dataset with ground truth for remote car pose estimation using satellite markers

Mr. Gustavo Gil, Giovanni Savino, Marco Pierini
University of Florence, Italy

Leading causes of PTW (Powered Two-Wheeler) crashes and near misses in urban areas are on the part of a failure or delayed prediction of the changing trajectories of other vehicles. Regrettably, misperception from both car drivers and motorcycle riders results in fatal or serious consequences for riders. Intelligent vehicles could provide early warning about possible collisions, helping to avoid the crash. There is evidence that stereo cameras can be used for estimating the heading angle of other vehicles, which is key to anticipate their imminent location, but there is limited heading ground truth data available in the public domain. Consequently, we employed a marker-based

technique for creating ground truth of car pose and create a dataset for computer vision benchmarking purposes. This dataset of a moving vehicle collected from a static mounted stereo camera is a simplification of a complex and dynamic reality, which serves as a test bed for car pose estimation algorithms. The dataset contains the accurate pose of the moving obstacle, and realistic imagery including texture-less and non-lambertian surfaces (e.g. reflectance and transparency).

V0090

Time: 14:15-14:30

Optimal frame-by-frame result combination strategy for OCR in video stream

Konstantin Bulatov, **Mr. Aleksander Lynchenko** and Valeriy Krivtsov
Smart Engines, Moscow, Russia

This paper describes the problem of combining classification results of multiple observations of one object. This task can be regarded as a particular case of a decision-making using a combination of experts votes with calculated weights. The accuracy of various methods of combining the classification results depending on different models of input data is investigated on the example of frame-by-frame character recognition in a video stream. Experimentally it is shown that the strategy of choosing a single most competent expert in case of input data without irrelevant observations has an advantage (in this case irrelevant means with character localization and segmentation errors). At the same time this work demonstrates the advantage of combining several most competent experts according to multiplication rule or voting if irrelevant samples are present in the input data.

V0114

Time: 14:30-14:45

Satellite Markers: a simple method for ground truth car pose on stereo video

Mr. Gustavo Gil, Giovanni Savino, Simone Piantini, Marco Pierini
University of Florence, Italy

Artificial prediction of future location of other cars in the context of advanced safety systems is a must. The remote estimation of car pose and particularly its heading angle is key to predict its future location. Stereo vision systems allow to get the 3D information of a scene. Ground truth in this specific context is associated with referential information about the depth, shape and orientation of the objects present in the traffic scene. Creating 3D ground truth is a measurement and data fusion task associated with the combination of different kinds of sensors. The novelty of this paper is the method to generate ground truth car pose only from video data. When the method is applied to stereo video, it also provides the extrinsic camera parameters for each camera at frame level which are key to quantify the performance of a stereo vision system when it is moving because the system is subjected to undesired vibrations and/or leaning. We developed a video post-processing technique which employs a common camera calibration tool for the 3D ground truth generation. In our case study, we focus in accurate car heading angle estimation of a moving car under realistic imagery. As outcomes, our satellite marker method provides accurate car pose at frame level, and the instantaneous spatial orientation for each camera at frame level.

V0132

Time: 14:45-15:00

Hierarchical VS Non-hierarchical Audio Indexation and Classification for Video Genres

Ms. Nouha Dammak and Yassine BenAyed
MIRACL Laboratory and Higher Institute of Computer Sciences and Communication Techniques, Tunisia

In this paper, Support Vector Machines (SVMs) are used for segmenting and indexing video genres based on only audio features extracted at block level, which has a prominent asset by capturing local temporal information. The main contribution of our study is to show the wide effect on the classification accuracies while using an hierarchical categorization structure based on Mel Frequency Cepstral Coefficients (MFCC) audio descriptor. In fact, the classification consists in three common video genres: sports videos, music clips and news scenes. The sub-classification may divide each genre into several multi-speaker and multi-dialect sub-genres. The validation of this approach was carried out on over 360 minutes of video span yielding a classification accuracy of over 99%.

V0134

Time: 15:00-15:15

A No-Reference Image and Video Visual Quality Metric Based on Machine Learning

Vladimir Frantc, **Mr. Viacheslav Voronin**, Evgenii Semenishchev, Maxim Minkin and Aliy Delov
Don State Technical University, Russian Federation

The paper presents a novel visual quality metric for lossy compressed video quality assessment. High degree of correlation with subjective estimations of quality is due to using of a convolutional neural network trained on a large amount of pairs video sequence-subjective quality score. We demonstrate how our predicted no-reference quality metric correlates with qualitative opinion in a human observer study. Results are shown on the EVVQ dataset with comparison existing approaches.

V0034

Time: 15:15-15:30

From Image Captioning to Video Summary using Deep Recurrent Networks and Unsupervised Segmentation

Morosanu Bogdan Andrei and **Lemnaru Camelia**
Technical University of Cluj-Napoca, Romania

Automatic captioning systems based on recurrent neural networks have been tremendously successful at providing realistic natural language captions for complex and varied image data. We explore methods for adapting existing models trained on large image caption data sets to a similar problem, that of summarising videos using natural language descriptions and frame selection. These architectures create internal high level representations of the input image that can be used to define probability distributions and distance metrics on these distributions. Specifically, we interpret each hidden unit inside a layer of the caption model as representing the un-normalised log probability of some unknown image feature of interest for the caption generation process. We can then apply well understood statistical divergences to express the difference between images and create an unsupervised segmentation of video frames, classifying consecutive images of low divergence as belonging to the same context, and those of high divergence as belonging to different contexts. To provide a final summary of the video, we provide a group of selected frames and a text description accompanying them, allowing a user to perform a quick exploration of large unlabeled video databases.

Session 5 < Machine vision and visualization >

Venue: Fleming's ROOM 6

Chair:

Time: 13:00pm-15:30pm

Note:
* The certificate of oral presentation will be awarded at the end of each session.
* For the Best Presentation of each session, it is encouraged to award it to student author prior. The best presentation will be announced at the end of session and awarded certificate over the dinner banquet.
*Please arrive the meeting room at least 10 Mins before session starts. To show the respect to other authors, especially to encourage the student authors, we strongly suggest you attend the whole session, the scheduled time for presentations might be changed due to unexpected situations, please come as early as you could.
* Session photos will be taken at the end of the session and update online

V0007

Time: 13:00-13:15

Wire Connector Classification with Machine Vision and a Novel Hybrid SVM

V.D. Chauhan, K.D. Joshi and **Prof. B.W. Surgenor**
Queen's University, Kingston, Canada

A machine vision-based system has been developed and tested that uses a novel hybrid Support Vector Machine (SVM) in a part inspection application with clear plastic wire connectors. The application required the system to differentiate between 4 different known styles of connectors plus one unknown style, for a total of 5 classes. The requirement to handle an unknown class is what necessitated the hybrid approach. The system was trained with the 4 known classes and tested with 5 classes (the 4 known plus the 1 unknown). The hybrid classification approach used two layers of SVMs: one layer was semi-supervised and the other layer was supervised. The semi-supervised SVM was a special case of unsupervised machine learning that classified test images as one of the 4 known classes (to accept) or as the unknown class (to reject). The supervised SVM classified test images as one of the 4 known classes and consequently would give false positives (FPs). Two methods were tested. The difference between the methods was that the order of the layers was switched. The method with the semi-supervised layer first gave an accuracy of 80% with 20% FPs. The method with the supervised layer first gave an accuracy of 98% with 0% FPs. Further work is being conducted to see if the hybrid approach works with other applications that have an unknown class requirement.

V0011

Time: 13:15-13:30

High-Resolution Hyperspectral Ground Mapping for Robotic Vision

Mr. Frank Neuhaus, Christian Fuchs, and Dietrich Paulus
University of Koblenz-Landau, Germany

Recently released hyperspectral cameras use large, mosaiced filter patterns to capture different ranges of the light's spectrum in each of the camera's pixels. Spectral information is sparse, as it is not fully available in each location. We propose an online method that avoids explicit demosaicing of camera images by fusing raw, unprocessed, hyperspectral camera frames inside an ego-centric ground surface map. It is represented as a multilayer heightmap data structure, whose geometry is estimated by combining a visual odometry system with either dense 3D reconstruction or 3D laser data.
We use a publicly available dataset to show that our approach is capable of constructing an accurate hyperspectral representation of the surface surrounding the vehicle. We show that in many cases our

approach increases spatial resolution over a demosaicing approach, while providing the same amount of spectral information.

V0113

Time: 13:30-13:45

Overview of Machine Vision Methods in X-ray Imaging and Microtomography

Dr. Alexey Buzmakov, Denis Zolotov, Marina Chukalina, Dmitry Nikolaev, Andrey Gladkov, Anastasia Ingacheva, Ivan Yakimchuk and Victor Asadchikov
Shubnikov Institute of Crystallography FSRC "Crystallography and Photonics" RAS, Moscow, Russia

Digital X-ray imaging became widely used in science, medicine, non-destructive testing. This allows using modern digital images analysis for automatic information extraction and interpretation. We give short review of scientific applications of machine vision in scientific X-ray imaging and microtomography, including image processing, feature detection and extraction, images compression to increase camera throughput, microtomography reconstruction, visualization and setup adjustment.

V0139

Time: 13:45-14:00

Fuzzy Classification for Strawberry Diseases-Infection Using Machine Vision and Soft-Computing Techniques

Mr. Hamit Altıparmak, Mohamad Al Shahadat, Ehsan Kiani, Kamil Dimililer
Department of Computer Engineering, Faculty of Engineering, Near East University, Nicosia, Cyprus

Robotic agriculture requires smart and doable techniques to substitute the human intelligence with machine intelligence. Strawberry is one of the important Mediterranean product and its productivity enhancement requires modern and machine-based methods. Whereas a human identifies the disease infected leaves by his eye, the machine should also be capable of vision-based disease identification. The objective of this paper is to practically verify the applicability of a new computer-vision method for discrimination between the healthy and disease infected strawberry leaves which does not require neural network or time consuming trainings. The proposed method was tested under outdoor lighting condition using a regular DLSR camera without any particular lens. Since the type and infection degree of disease is approximated a human brain a fuzzy decision maker classifies the leaves over the images captured on-site having the same properties of human vision. Optimizing the fuzzy parameters for a typical strawberry production area at a summer mid-day in Cyprus produced 96% accuracy for segmented iron deficiency and 93% accuracy for segmented using a typical human instant classification approximation as the benchmark holding higher accuracy than a human eye identifier. The fuzzy-base classifier provides approximate result for decision making on the leaf status as if it is healthy or not.

V0051

Time: 14:00-14:15

Drawing a baseline in Aesthetic Quality Assessment

Mr. Fernando Rubio Perona, M. Julia Flores Gallego and Jose Miguel Puerta Callejón
SIMD Lab, I3A, UCLM, Albacete, Spain

Aesthetic classification of images is an inherently subjective task. There does not exist a validated

collection of images/photographs labeled as having good or bad quality from experts. Nowadays, the closest approximation to that is to use databases of photos where a group of users rate each image. Hence, there is not a unique good/bad label but a rating distribution given by users *voting*.
Due to this peculiarity, it is not possible to state the problem of binary aesthetic supervised classification in such a direct mode as other Computer Vision tasks. Recent literature follows an approach where researchers utilize the average rates from the users for each image, and they establish an arbitrary threshold to determine their class or label. In this way, images above the threshold are considered of good quality, while images below the threshold are seen as bad quality.
This paper analyzes current literature, and it reviews those attributes able to represent an image, differentiating into three families: specific, general and deep features. Among those which have been proved more competitive, we have selected a representative subset, being our main goal to establish a clear experimental framework. Finally, once features were selected, we have used them for the full AVA dataset. We have to remark that to perform validation we report not only accuracy values, which is not that informative in this case, but also, metrics able to evaluate classification power within imbalanced datasets. We have conducted a series of experiments so that distinct well-known classifiers are learned from data. Like that, this paper provides what we could consider valuable and valid *baseline* results for the given problem.

V0066

Time: 14:15-14:30

Computer vision system: a tool fpr evaluating the quality of wheat in a grain tank

Uryi Igorevish Minkin, Aleksei Vladimirovich Panchenko, **Mr. Aleksandr Yurievich Shkanaev**, Ivan Andeevich Konovalenko, Dmitry Nikolaevich Putintsev, Rinat Nailevish Sadekov
Cognitive Technologies LTD, Russian Federation

The paper describes a technology that allows for automatizing the process of evaluating the grain quality in a grain tank of a combine harvester. Special recognition algorithm analyzes photographic images taken by the camera, and that provides automatic estimates of the total mass fraction of broken grains and the presence of non-grains. The paper also presents the operating details of the tank prototype as well as it defines the accuracy of the algorithms designed.

V0017

Time: 14:30-14:45

A Study on Low-cost, High-accuracy and Real-time Stereo Vision Algorithms for UAV Power Line Inspection

Hongyu Wang, **Mr. Baomin Zhang**, Xun Zhao, Cong Li and Cunyue Lu
Shanghai Jiao Tong University, China

Conventional stereo vision algorithms suffer from high levels of hardware resource utilization due to algorithm complexity, or poor levels of accuracy caused by inadequacies in the matching algorithm. To address these issues, we have proposed a stereo range-finding technique that produces an excellent balance between cost, matching accuracy and real-time performance, for power line inspection using UAV. This was achieved through the introduction of a special image preprocessing algorithm and a weighted local stereo matching algorithm, as well as the design of a corresponding hardware architecture. Stereo vision systems based on this technique have a lower level of resource usage and also a higher level of matching accuracy following hardware acceleration. To validate the effectiveness of our technique, a stereo vision system based on our improved algorithms were implemented using the Spartan 6 FPGA. In comparative experiments, it was shown that the system using the improved algorithms outperformed the system based on the un-improved algorithms, in terms of resource utilization and matching accuracy. In particular, Block RAM usage was reduced

by 19%, and the improved system was also able to output range-finding data in real time.

V0065

Time: 14:45-15:00

Aerial images visual localization on a vector map using color-texture segmentation

Ms. Irina Kunina, Lev Teplyakov, Andrey Gladkov, Timur Khanipov and Dmitry Nikolaev
IITP RAS, Russia

In this paper we study the problem of combining UAV obtained optical data and a coastal vector map in absence of satellite navigation data. The method is based on presenting the territory as a set of segments produced by color-texture image segmentation. We then find such geometric transform which gives the best match between these segments and land and water areas of the georeferenced vector map. We calculate transform consisting of an arbitrary shift relatively to the vector map and bound rotation and scaling. These parameters are estimated using the RANSAC algorithm which matches the segments contours and the contours of land and water areas of the vector map. To implement this matching we suggest computing shape descriptors robust to rotation and scaling. We performed numerical experiments demonstrating the practical applicability of the proposed method.

V0135

Time: 15:00-15:15

Using virtual environment for autonomous vehicle algorithm validation

Mr. Aleksandrs Levinskis
Institute of Electronics and Comuter Science (EDI), Latvia

This paper describes possible use of modern game engine for validating and proving the concept of algorithm design. As the result simple visual odometry algorithm will be provided to show the concept and go over all workflow stages. Some of stages will involve using of Kalman filter in such a way that it will estimate optical flow velocity as well as position of moving camera located at vehicle body. In particular Unreal Engine 4 game engine will be used for generating optical flow patterns and ground truth path. For optical flow determination Horn and Schunck method will be applied. As the result, it will be shown that such method can estimate position of the camera attached to vehicle with certain displacement error respect to ground truth depending on optical flow pattern. For displacement rate RMS error is calculating between estimated and actual position.

Session 6 < Image analysis and imaging system >

Venue: Fleming's ROOM 9

Chair: Prof. Henryk Palus, Silesian University of Technology, Gliwice, Poland

Time: 15:45pm-18:30pm

Note:

* The certificate of oral presentation will be awarded at the end of each session.
* For the Best Presentation of each session, it is encouraged to award it to student author prior. The best presentation will be announced at the end of session and awarded certificate over the dinner banquet.
*Please arrive the meeting room at least 10 Mins before session starts. To show the respect to other authors, especially to encourage the student authors, we strongly suggest you attend the whole session, the scheduled time for presentations might be changed due to unexpected situations, please come as early as you could.
* Session photos will be taken at the end of the session and update online

Plenary Speech

Time: 15:45-16:15

Toward a perceptual image quality assessment of color quantized images

Mariusz Frackiewicz and **Prof. Henryk Palus**
Silesian University of Technology, Gliwice, Poland

Color image quantization is an important operation in the _eld of color image processing. In this paper, we consider new perceptual image quality metrics for assessment of quantized images. These types of metrics, e.g. DSCSI, MDSIs, MDSIm and HPSI achieve the highest correlation coe_cients with MOS during tests on the six publicly available image databases. Research was limited to images distorted by two types of compression: JPG and JPG2K. Statistical analysis of correlation coe_cients based on the Friedman test and post-hoc procedures showed that the di_erences between the four new perceptual metrics are not statistically signi_cant.

V0072

Time: 16:15-16:30

Pixel-wise Deblurring Imaging System Based on Active Vision for Structural Health Monitoring at a Speed of 100 km/h

Dr. Tomohiko Hayakawa, Yushi Moko, Kenta Morishita, Masatoshi Ishikawa
The University of Tokyo, Japan

In this paper, we propose a pixel-wise deblurring imaging (PDI) system based on active vision for compensation of the blur caused by high-speed one-dimensional motion between a camera and a target. The optical axis is controlled by back-and-forth motion of a galvanometer mirror to compensate the motion. High-spatial-resolution image captured by our system in high-speed motion is useful for efficient and precise visual inspection, such as visually judging abnormal parts of a tunnel surface to prevent accidents; hence, we applied the PDI system for structural health monitoring. By mounting the system onto a vehicle in a tunnel, we confirmed significant improvement in image quality for submillimeter black-and-white stripes and real tunnel-surface cracks at a speed of 100 km/h.

V0107

Time: 16:30-16:45

Blur kernel estimation with algebraic tomography technique and intensity profiles of object boundaries

Ms. Anastasia Ingacheva, Marina Chukalina, Timur Khanipov and Dmitry Nikolaev
IITP, Russia

Motion blur caused by camera vibration is a common source of degradation in photographs. In this paper we study the problem of finding the point spread function (PSF) of a blurred image using the tomography technique. The PSF reconstruction result strongly depends on the particular tomography technique used. We present a tomography algorithm with regularization adapted specifically for this task. We use the algebraic reconstruction technique (ART algorithm) as the starting algorithm and introduce regularization. We use the conjugate gradient method for numerical implementation of the proposed approach. The algorithm is tested using a dataset which contains 9 kernels extracted from real photographs by the Adobe corporation where the point spread function is known. We also investigate influence of noise on the quality of image reconstruction and investigate how the number of projections influence the magnitude change of the reconstruction error.

V0056

Time: 16:45-17:00

Training a whole-book LSTM-based recognizer with an optimal training set

Assist. Prof. Mohammad Reza Soheili, Mohammad Reza Yousefi, Ehsanollah Kabir and Didier Stricker
Kharazmi University, Iran

Despite the recent progress in OCR technologies, whole-book recognition, is still a challenging task, in particular in case of old and historical books, that the unknown font faces or low quality of paper and print contributes to the challenge. Therefore, pre-trained recognizers and generic methods do not usually perform up to required standards, and usually the performance degrades for larger scale recognition tasks, such as of a book. Such reportedly low error-rate methods turn out to require a great deal of manual correction. Generally, such methodologies do not make effective use of concepts such redundancy in whole-book recognition. In this work, we propose to train Long Short Term Memory (LSTM) networks on a minimal training set obtained from the book to be recognized. We show that

clustering all the sub-words in the book, and using the sub-word cluster centers as the training set for the LSTM network, we can train models that outperform any identical network that is trained with randomly selected pages of the book. In our experiments, we also show that although the sub-word cluster centers are equivalent to about 8 pages of text for a 101-page book, a LSTM network trained on such a set performs competitively compared to an identical network that is trained on a set of 60 randomly selected pages of the book.

V0082

Time: 17: 00-17:15

3-D shape recovery from Image focus using Gray Level Co-Occurance Matrix

Mr. Fahad Mahmood, Umair Munir, Fahad Mehmood, Javaid Iqbal
NUST, Pakistan

Recovering a precise and accurate 3-D shape of the target object utilizing robust 3-D shape recovery algorithm is an ultimate objective of computer vision community. Focus measure algorithm plays an important role in this architecture which convert the color values of each pixel of the acquired 2-D image dataset into corresponding focus values. After convolving the focus measure filter with the input 2-D image dataset, a 3-D shape recovery approach is applied which will recover the depth map. In this document, we are concerned with proposing Gray Level Co-occurrence Matrix along with its statistical features for computing the focus information of the image dataset. The Gray Level Co-occurrence Matrix quantifies the texture present in the image using statistical features and then applies joint probability distributive function of the gray level pairs of the input image. Finally, we quantify the focus value of the input image using Gaussian Mixture Model. Due to its little computational complexity, sharp focus measure curve, robust to random noise sources and accuracy, it is considered as superior alternative to most of recently proposed 3-D shape recovery approaches. This algorithm is deeply investigated on real image sequences and synthetic image dataset. The efficiency of the proposed scheme is also compared with the state of art 3-D shape recovery approaches. Finally, by means of two global statistical measures, root mean square error and correlation, we claim that this approach –in spite of simplicity- generates accurate results.

V0103

Time: 17: 15-17:30

Document Localization Algorithms Based On Feature Points And Straight Lines

Ms. Natalya Skoryukina, Julia Shemiakina, Vladimir Arlazarov, Igor Faradjev
Smart Engines, Russia

The important part of the system of a planar rectangular object analysis is the localization: the estimation of projective transform from template image of an object to its photograph. The system also includes such subsystems as the selection and recognition of text fields, the usage of contexts etc. In this paper three localization algorithms are described. All algorithms use feature points and two of them also analyze near-horizontal and near- vertical lines on the photograph. The algorithms and their combinations are tested on a dataset of real document photographs. Also the method of localization quality estimation is proposed that allows configuring the localization subsystem independently of the other subsystems quality.

V0111

Time: 17: 30-17:45

Classification of Time-Series Images Using Deep Convolutional Neural Networks

Assoc. Prof. Nima Hatami, Yann Gavet, Johan Debayle
Ecole Nationale Supérieure des Mines de Saint-Etienne

Convolutional Neural Networks (CNN) has achieved a great success in image recognition task by automatically learning a hierarchical feature representation from raw data. While the majority of Time-Series Classi_cation (TSC) literature is focused on 1D signals, this paper uses Recurrence Plots (RP) to transform time-series into 2D texture images and then take advantage of the deep CNN classi_er. Image representation of time-series introduces di_erent feature types that are not available for 1D signals, and therefore TSC can be treated as texture image recognition task. CNN model also allows learning di_erent levels of representations together with a classi_er, jointly and automatically. Therefore, using RP and CNN in a uni_ed framework is expected to boost the recognition rate of TSC. Experimental results on the UCR time-series classi_cation archive demonstrate competitive accuracy of the proposed approach, compared not only to the existing deep architectures, but also to the state-of-the art TSC algorithms.

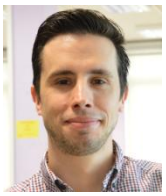
V0099A

Time: 17:45-18:00

3D image analysis in medical applications

Dr. I. Bukreeva, G. Campi, M. Fratini, L. Massimi & A. Cedola
Nanotec-CNR, c/o Phys. Dep. at 'Sapienza' University, Rome, Italy

Recent progress in X-ray technique and imaging clearly highlights the need for development new approaches to 3D visualization, image processing, pattern recognition and, as a final goal, distinct 3D qualitative and quantitative characterization of objects. In particular, we present high (micrometer and sub-micrometer level) resolution 3D images of murine spinal cord (SP) preserving all richness of micro-details in neuronal and vascular architecture [1-4]. High quality of images enables easy 3D qualitative and quantitative analysis of organ on the level inaccessible until recently. In particular we developed and applied numerical method facilitating a quantitative 3D spatial statistical analysis of the neuronal structure of SP. Proposed method can be considered as additional contribution to investigation of neurodegenerative diseases and therapies.



V0102

Time: 18: 00-18:15

Spatio-thermal depth correction of RGB-D sensors based on Gaussian Processes in real-time

Mr. Christoph Heindl, Thomas Pönitz, Gernot Stubl, Andreas Pichler, and Josef Scharinger
4407 Steyr Gleink, Austria

Commodity RGB-D sensors capture color images along with dense pixel-wise depth information in real-time. Typical RGB-D sensors are provided with a factory calibration and exhibit erratic depth readings due to coarse calibration values, ageing and thermal influence effects. This limits their applicability in computer vision and robotics. We propose a novel method to accurately calibrate depth considering spatial and thermal influences jointly. Our work is based on Gaussian Process Regression in a four dimensional Cartesian and thermal domain. We propose to leverage modern GPUs for dense depth map correction in real-time. For reproducibility we make our dataset and source code publicly available.

V0040

Time: 18: 15-18:30

Copy Move Image Forgery Detection Based on Polar Fourier Representation

Yitian Wang and **Prof. Sei-ichiro Kamata**
Waseda University, Japan

With the rapid development of multimedia technology, it's easy for someone to obtain an image and edit it according to their own preferences or some ulterior purpose. Copy–Move is a common type of digital image forgery where a part of the original image is copied and pasted at another position in the same image. In this paper, we propose an efficient methodology for enhancing block matching based on Copy–Move forgery detection. The main contribution of this work is the utilization of polar representation to get the representative features for each block. The main feature used in this paper is the frequency of each block based on Fourier transform. The experimental results show the efficiency of the proposed method for detecting copy-move regions, even when the copied region has undergone severe image manipulations such as rotation, scaling, Gaussian blurring, brightness modification, JPEG compression and noise addition.

Session 7 < Image processing techniques and methods >

Venue: Fleming's ROOM 10

Chair:

Time: 15:45pm-18:15pm

Note:

* The certificate of oral presentation will be awarded at the end of each session.
* For the Best Presentation of each session, it is encouraged to award it to student author prior. The best presentation will be announced at the end of session and awarded certificate over the dinner banquet.
*Please arrive the meeting room at least 10 Mins before session starts. To show the respect to other authors, especially to encourage the student authors, we strongly suggest you attend the whole session, the scheduled time for presentations might be changed due to unexpected situations, please come as early as you could.
* Session photos will be taken at the end of the session and update online

V0025

Time: 15:45-16:00

Sum of Top-Hat Transform Based Algorithm for Vessel enhancement in MRA images

Dr. Ouazaa Hibet-Allah, Jlassi Hajer and Hamrouni Kamel
LR-SITI-ENIT, Tunisia

The Magnetic Resonance Angiography (MRA) is rich with information's. But, they suffer from poor contrast, illumination and noise. Thus, it is required to enhance the images. But, these significant information can be lost if improper techniques are applied. Therefore, in this paper, we propose a new method of enhancement. We applied firstly the CLAHE method to increase the contrast of the image. Then, we applied the sum of Top-Hat Transform to increase the brightness of vessels. It is performed with the structuring element oriented in different angles. The methodology is tested and evaluated on the publicly available database BRAINIX. And, we used the measurement methods MSE (Mean Square Error), PSNR (Peak Signal to Noise Ratio) and SNR (Signal to Noise Ratio) for the evaluation. The results demonstrate that the proposed method could efficiently enhance the image details and is comparable with state of the art algorithms. Hence, the proposed method could be broadly used in various applications.

V0026

Time: 16:00-16:15

Image Quality Assessment for Selfies with and without Super Resolution

Aya Kubota and **Prof. Seiichi Gohshi**
Seiichi Gohshi, Kogakin University, Japan

With the advent of cellphone cameras, in particular, on smartphones, many people now take photos of themselves alone and with others in the frame; such photos are popularly known as "selfies". Most smartphones are equipped with two cameras: the front-facing and rear cameras. The camera located on the back of the smartphone is referred to as the "out-camera," whereas the one located on the front of the smartphone is called the "in-camera." In-cameras are mainly used for selfies. Some smartphones feature high-resolution cameras. However, the original image quality cannot be obtained because smartphone cameras often have low-performance lenses. Super resolution (SR) is one of the recent technological advancements that has increased image resolution. We developed a new SR technology that can be processed on smartphones. Smartphones with new SR technology are currently available in the market have already registered sales. However, the effective use of new SR technology has not yet been verified. Comparing the image quality with and without SR on smartphone display is necessary to confirm the usefulness of this new technology. Methods that are based on objective and subjective assessments are required to quantitatively measure image

quality. It is known that the typical object assessment value, such as Peak Signal to Noise Ratio (PSNR), does not go together with how we feel when we assess image/video. When digital broadcast started, the standard was determined using subjective assessment. Although subjective assessment usually comes at high cost because of personnel expenses for observers, the results are highly reproducible when they are conducted under right conditions and statistical analysis. In this study, the subjective assessment results for selfie images are reported.

V0031

Time: 16:15-16:30

Multiview 3D sensing and analysis for high quality point cloud reconstruction.
Mr. Andrej Satnik, Richard Orjesek, Ebroul Izquierdo
Queen Mary University of London, United Kingdom

Multiview 3D reconstruction techniques enable digital reconstruction of 3D objects from the real world by fusing different viewpoints of the same object into a single 3D representation. This process is by no means trivial and the acquisition of high quality point cloud representations of dynamic 3D objects is still an open problem. In this paper, an approach for high fidelity 3D point cloud generation using low cost 3D sensing hardware is presented. The proposed approach runs in an efficient low-cost hardware setting based on several Kinect v2 scanners connected to a single PC. It performs autocalibration and runs in real-time exploiting an efficient composition of several filtering methods including Radius Outlier Removal (ROR), Weighted Median filter (WM) and Weighted Inter-Frame Average filtering (WIFA). The performance of the proposed method has been demonstrated through efficient acquisition of dense 3D point clouds of moving objects.

V0043

Time: 16:30-16:45

Singular spectrum decomposition of Bouligand-Minkowski fractal descriptors: an application to the classification of texture images

Prof. Joao Batista Florindo
University of Campinas, Brazil

This work proposes the use of Singular Spectrum Analysis (SSA) for the classification of texture images, more specifically, to enhance the performance of the Bouligand-Minkowski fractal descriptors in this task. Fractal descriptors are known to be a powerful approach to model and particularly identify complex patterns in natural images. Nevertheless, the multiscale analysis involved in those descriptors makes them highly correlated. Although other attempts to address this point was proposed in the literature, none of them investigated the relation between the fractal correlation and the well-established analysis employed in time series. And SSA is one of the most powerful techniques for this purpose. The proposed method was employed for the classification of benchmark texture images and the results were compared with other state-of-the-art classifiers, confirming the potential of this analysis in image classification.

V0049

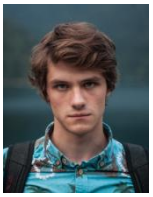
Time: 16:45-17:00

Non parametric adaptative JPEG fragments carving

Ms. Sabrina Amrouche and Dalila Salamani
Geneva university, Switzerland

The most challenging JPEG recovery tasks arise when the file header is missing. In this paper we propose to use a two layer machine learning model to restore headerless JPEG images. We first

build a classifier able to identify the structural properties of the images/fragments and then use an AutoEncoder (AE) to learn the fragment features for the header prediction. We define a JPEG universal header and the remaining free image parameters (Height, Width) are predicted with a Gradient Boosting Classifier. Our approach resulted in 90% accuracy using the manually defined features and 78% accuracy using the AE features



V0074

Time: 17:00-17:15

Perception-oriented fusion of multi-sensor imagery: visible, IR and SAR

Mr. D. Sidorchuk and V. Volkov
Institute for Information Transmission Problems, Russia

This paper addresses the problem of image fusion of optical (visible and thermal domain) data and radar data for the purpose of visualization. These types of images typically contain a lot of complimentary information, and their joint visualization can be useful and more convenient for human user than a set of individual images. To solve the image fusion problem we propose a novel algorithm that utilizes some peculiarities of human color perception and based on the grey-scale structural visualization. Benefits of presented algorithm are exemplified by satellite imagery.

V0104

Time: 17:15-17:30

Fast Words Boundaries Localization in Text Fields for Low Quality Document Images

Mr. Dmitry Ilin, Dmitriy Novikov, Dmitry Polevoy and Dmitry Nikolaev
Smart Engines Ltd, Moscow

The paper examines the problem of word boundaries precise localization in document text zones. Document processing on a mobile device consists of document localization, perspective correction, localization of individual fields, finding words in separate zones, segmentation and recognition. While capturing an image with a mobile digital camera under uncontrolled capturing conditions, digital noise, perspective distortions or glares may occur. Further document processing gets complicated because of its specifics: layout elements, complex background, static text, document security elements, variety of text fonts. However, the problem of word boundaries localization has to be solved at run-time on mobile CPU with limited computing capabilities under specified restrictions. At the moment, there are several groups of methods optimized for different conditions. Methods for the scanned printed text are quick but limited only for images of high quality. Methods for text in the wild have an excessively high computational complexity, thus, are hardly suitable for running on mobile devices as part of the mobile document recognition system. The method presented in this paper solves a more specialized problem than the task of finding text on natural images. It uses local features, a sliding window and a lightweight neural network in order to achieve an optimal algorithm speed-precision ratio. The duration of the algorithm is 12 ms per field running on an ARM processor of a mobile device. The error rate for boundaries localization on a test sample of 8000 fields is 0.3

V0115

Time: 17:30-17:45

The evaluation of correction algorithms of intensity nonuniformity in breast MRI images: a phantom study

Damian Borys, Wojciech Serafin, Kamil Gorczewski, Marek Kijonka, **Assoc. Prof. Mariusz**

Frackiewicz, Henryk Palus
Silesian University of Technology, Gliwice, Poland

The aim of this work was to test the most popular and essential algorithms of the intensity nonuniformity correction of the breast MRI imaging. Evaluated methods were: N3, N3FCM, N4, Nonparametric, and SPM. For testing purposes, a uniform phantom object was used to obtain test images. To quantify the results, two measures were used: integral uniformity and standard deviation. In the result, two methods obtained the lowest values in these measures: N3FCM and N4, however, for the second method visually phantom was the most uniform after correction.

V0133

Time: 17:45-18:00

Reducing Noise Component On Medical Images

Evgeny Semenishchev, **Mr. Vyacheslav Voronin** and Vladimir Dub
Don State Technical University, Russian Federation

Medical visualization and analysis of medical data is an actual direction. Medical images are used in microbiology, genetics, roentgenology, oncology, surgery, ophthalmology, etc. Initial data processing is a major step towards obtaining a good diagnostic result. The paper considers the approach allows an image filtering with preservation of objects borders. The algorithm proposed in this paper is based on sequential data processing. At the first stage, local areas are determined, for this purpose the method of threshold processing, as well as the classical ICI algorithm, is applied. The second stage uses a method based on based on two criteria, namely, L2 norm and the first order square difference. To preserve the boundaries of objects, we will process the transition boundary and local neighborhood the filtering algorithm with a fixed-coefficient. For example, reconstructed images of CT, x-ray, and microbiological studies are shown. The test images show the effectiveness of the proposed algorithm. This shows the applicability of analysis many medical imaging applications.

V0101

Time: 18:00-18:15

The Method for Homography Estimation between Two Planes based on Lines and Points

Ms. Julia Shemiakina, Alexander Zhukovsky, Dmitry Nikolaev
Smart Engines, Moscow

The paper considers the problem of estimating a transform connecting two images of one plane object. The method based on RANSAC is proposed for calculating the parameters of projective transform which uses points and lines correspondences simultaneously. A series of experiments was performed on synthesized data. Presented results show that the algorithm convergence rate is significantly higher when actual lines are used instead of points of lines intersection. When using both lines and feature points it is shown that the convergence rate does not depend on the ratio between lines and feature points in the input dataset.

Session 8< Information system and image application >

Venue: Fleming's ROOM 6

Chair:

Time: 15:45pm-18:30pm

Note:

* The certificate of oral presentation will be awarded at the end of each session.
* For the Best Presentation of each session, it is encouraged to award it to student author prior. The best presentation will be announced at the end of session and awarded certificate over the dinner banquet.
*Please arrive the meeting room at least 10 Mins before session starts. To show the respect to other authors, especially to encourage the student authors, we strongly suggest you attend the whole session, the scheduled time for presentations might be changed due to unexpected situations, please come as early as you could.
* Session photos will be taken at the end of the session and update online

V0137

Time: 15:45-16:00

Analysis of Computer Images in the Presence of Metals

Alexey Buzmakov, Anastasia Ingacheva, Victor Prun, Dmitry Nikolaev, **Dr. Marina Chukalina**, Claudio Ferrero and Victor Asadchikov
FSRC "Crystallography and Photonics" RAS, Russia

Artefacts caused by intensely absorbing inclusions are encountered in computed tomography via polychromatic scanning and may obscure or simulate pathologies in medical applications. To improve the quality of reconstruction if high-Z inclusions in presence, previously we proposed and tested with synthetic data an iterative technique with soft penalty mimicking linear inequalities on the photon-starved rays. This note reports a test at the tomographic laboratory set-up at the Institute of Crystallography FSRC "Crystallography and Photonics" RAS in which tomographic scans were successfully made of temporary tooth without inclusion and with Pb inclusion.

V0048

Time: 16:00-16:15

Triadic Split-Merge Sampler

Mr. Anne van Rossum, Johan Dubbeldam, Hai Xiang Lin and H.Jaap van den Herik
Delft University of Technology, Crownstone

In machine vision typical heuristic methods to extract parameterized objects out of raw data points are the Hough transform and RANSAC. Bayesian models carry the promise to optimally extract such parameterized objects given a correct definition of the model and the type of noise at hand. A category of solvers for Bayesian models are Markov chain Monte Carlo methods. Naive implementations ofMCMCmethods suffer from slow convergence in machine vision due to the complexity of the parameter space. Towards this blocked Gibbs and split-merge samplers have been developed that assign multiple data points to clusters at once. In this paper we introduce a new split-merge sampler, the triadic split-merge sampler, that perform steps between two and three randomly chosen clusters. This has two advantages. First, it reduces the asymmetry between the split and merge steps. Second, it is able to propose a new cluster that is composed out of data points from two different clusters. Both advantages speed up convergence which we demonstrate on a line extraction problem. We show that the triadic split-merge sampler outperforms the conventional split-merge sampler. Although this new MCMC sampler is demonstrated in this machine vision context, its application extend to the very general domain of statistical inference.

V0053

Time: 16:15-16:30

Face landmark point tracking using LK pyramid optical flow

Dr. Gang Zhang, Sikan Tang, Jiaquan Li
Shenyang University of Technology, China

LK pyramid optical flow is an effective method to implement object tracking in a video. It is used for face landmark point tracking in a video in the paper. The landmark points, *i.e.* outer corner of left eye, inner corner of left eye, inner corner of right eye, outer corner of right eye, tip of a nose, left corner of mouth, right corner of mouth, are considered. It is in the first frame that the landmark points are marked by hand. For subsequent frames, performance of tracking is analyzed. Two kinds of conditions are considered, *i.e.* single factors such as normalized case, pose variation and slowly moving, expression variation, illumination variation, occlusion, front face and rapidly moving, pose face and rapidly moving, and combination of the factors such as pose and illumination variation, pose and expression variation, pose variation and occlusion, illumination and expression variation, expression variation and occlusion. Global measures and local ones are introduced to evaluate performance of tracking under different factors or combination of the factors. The global measures contain the number of images aligned successfully, average alignment error, the number of images aligned before failure, and the local ones contain the number of images aligned successfully for components of a face, average alignment error for the components. To testify performance of tracking for face landmark points under different cases, tests are carried out for image sequences gathered by us. Results show that the LK pyramid optical flow method can implement face landmark point tracking under normalized case, expression variation, illumination variation which does not affect facial details, pose variation, and that different factors or combination of the factors have different effect on performance of alignment for different landmark points.

V0058

Time: 16:30-16:45

Ontology Based Decision System for Breast Cancer Diagnosis

Dr. Soumaya Trabelsi Ben Ameur, Dorra Sellami, Florence Cloppet and Laurent Wendling
ENIS-Tunisia/Paris Descartes-France

In this paper, we focus on analysis and diagnosis of breast masses inspired by expert concepts and rules. Accordingly, a Bag of Words is built based on the ontology of breast cancer diagnosis, accurately described in the Breast Imaging Reporting and Data System. To fill the gap between low level knowledge and expert concepts, a semantic annotation is developed using a machine learning tool. Then, breast masses are classified into benign or malignant according to expert rules implicitly modeled with a set of classifiers (KNN, ANN, SVM and Decision Tree). This semantic context of analysis offers a frame where we can include external factors and other meta-knowledge such as patient risk factors as well as exploiting more than one modality. Based on MRI and DECEDM modalities, our developed system leads a recognition rate of 99.7% with Decision Tree where an improvement of 24.7 % is obtained owing to semantic analysis.

V0073

Time: 16:45-17:00

Method of Estimation of Scanning System Quality

Eugene Larkin, Alexander Privalov, Vladislav Kotov, **Assoc. Prof. Natalya Kotova**
Tula State University (Tula, Russia)

Estimation of scanner parameters is an important part in developing electronic document management system. This paper suggests considering the scanner as a system that contains two main channels: a photoelectric conversion channel and a channel for measuring spatial coordinates of objects. Although both of channels consist of the same elements, the testing of their parameters should be executed separately. The special structure of the two-dimensional reference signal is offered for this purpose. In this structure, the fields for testing various parameters of the scanner are sp atially separated. Characteristics of the scanner are associated with the loss of information when a document is digitized. The methods to test grayscale transmitting ability, resolution and aberrations level are offered.

V0080

Time: 17:00-17:15

Laser Projection Positioning of Spatial Contour Curves via a Galvanometric Scanner

Mr. Junchao Tu, Liyan Zhang
Nanjing University of Aeronautics and Astronautics, China

The technology of laser projection positioning is widely applied in advanced manufacturing fields (e.g. composite plying, parts location and installation). In order to use it better, a laser projection positioning (LPP) system is designed and implemented. Firstly, the LPP system is built by a laser galvanometric scanning (LGS) system and a binocular vision system. Applying Single-hidden Layer Feed-forward Neural Network (SLFN), the system model is constructed next. Secondly, the LGS system and the binocular system, which are respectively independent, are integrated through a data-driven calibration method based on extreme learning machine (ELM) algorithm. Finally, a projection positioning method is proposed within the framework of the calibrated SLFN system model. A well-designed experiment is conducted to verify the viability and effectiveness of the proposed system. In addition, the accuracy of projection positioning are evaluated to show that the LPP system can achieves the good localization effect.

V0105

Time: 17:15-17:30

Generation method of synthetic training data for mobile OCR system

Ms. Yulia S. Chernyshova, Alexander V. Gayer, Alexander V. Sheshkus
NUST “MISIS”, Russia

This paper addresses one of the fundamental problems of machine learning - training data acquiring. Obtaining enough natural training data is rather difficult and expensive. In last years usage of synthetic images has become more beneficial as it allows to save human time and also to provide a huge number of images which otherwise would be difficult to obtain. However, for successful learning on artificial dataset one should try to reduce the gap between natural and synthetic data distributions. In this paper we describe an algorithm which allows to create artificial training datasets for OCR systems using russian passport as a case study.

V0106

Time: 17:30-17:45

Automatic White Blood Cell Classification using Pre-trained Deep Learning Models: ResNet and Inception

Mehdi Habibzadeh, **Ms. Mahboobeh Jannesari**, Zahra Rezaei, Mehdi Totonchi, and Hossein aharvand
University of Tehran & Iran

This works gives an account of evaluation of white blood cell differential counts via computer aided diagnosis (CAD) system and hematology rules. Leukocytes, also called white blood cells (WBCs) play main role of the immune system. Leukocyte is responsible for phagocytosis and immunity and therefore in defense against infection involving the fatal diseases incidence and mortality related issues. Admittedly, microscopic examination of blood samples is a time consuming, expensive and error-prone task. A manual diagnosis would search for specific Leukocytes and number abnormalities in the blood slides while complete blood count (CBC) examination is performed. Complications may arise from the large number of varying samples including different types of Leukocytes, related sub-types and concentration in blood, which makes the analysis prone to human error. This process can be automated by computerized techniques which are more reliable and economical. In essence, we seek to determine a fast, accurate mechanism for classification and gather information about distribution of white blood evidences which may help to diagnose the degree of any abnormalities during CBC test. In this work, we consider the problem of pre-processing and supervised classification of white blood cells into their four primary types including Neutrophils, Eosinophils, Lymphocytes, and Monocytes using a consecutive proposed deep learning framework. For first step, this research proposes three consecutive pre-processing calculations namely are color distortion; bounding box distortion (crop) and image flipping mirroring. In second phase, white blood cell recognition performed with hierarchy topological feature extraction using Inception and ResNet architectures. Finally, the results obtained from the preliminary analysis of cell classification with (11200) training samples and 1244 white blood cells evaluation data set are presented in confusion matrices and interpreted using accuracy rate, and false positive with the classification framework being validated with experiments conducted on poor quality blood images sized 320 × 240 pixels. The deferential outcomes in the challenging cell detection task, as shown in result section, indicate that there is a significant achievement in using Inception and ResNet architecture with proposed settings. Our framework detects on average 100% of the four main white blood cell types using ResNet V1 50 while also alternative promising result with 99.84% and 99.46% accuracy rate obtained with ResNet V1 152 and ResNet 101, respectively with 3000 epochs and fine-tuning all layers. Further statistical confusion matrix tests revealed that this work achieved 1, 0.9979, 0.9989 sensitivity values when area under the curve (AUC) scores above 1, 0.9992, 0.9833 on three proposed techniques. In addition, current work shows negligible and small false negative 0, 2, 1 and substantial false positive with 0, 0, 5 values in Leukocytes detection.

V0122

Time: 17:45-18:00

Convolutional neural network with transfer learning for rice type classification

Mr. Vaibhav Amit Patel and Manjunath V. Joshi
DA-IICT, India

Presently, rice type is identified manually by humans, which is time consuming and error prone. Therefore, there is a need to do this by machine which makes it faster with greater accuracy. This paper proposes a deep learning based method for classification of rice types. We propose two methods to classify the rice types. In the first method, we train a deep convolutional neural network (CNN) using the given segmented rice images. In the second method, we train a combination of a pretrained VGG16 network and the proposed method, while using transfer learning in which the weights of a pretrained network are used to achieve better accuracy. Our approach can also be used for classification of rice grain as broken or fine. We train a 5-class model for classifying rice types using 4000 training images and another 2-class model for the classification of broken and normal rice using 1600 training images. We observe that despite having distinct rice images, our architecture, pretrained on ImageNet data boosts classification accuracy significantly.

V0125

Time: 18:00-18:15

Prospects of improving the self-driving car development pipeline: transfer of algorithms from virtual to physical environment

Mr. Nauris Dorbe, Ingars Ribners, Krisjanis Nesenbergs
Institute of Electronics and Comuter Science (EDI), Latvia

Problem of transferring and testing self-driving algorithms developed in virtual environment to a physical environment is explored by transferring a Convolutional Neural Network based self-driving car steering algorithm from virtual environment to physical RC card based environment for validation and testing as a step on the way for full scale self-driving car tests. In the process a novel approach for synthetic training data generation from single camera is developed, thus reducing the real world physical requirements for the algorithm and demonstrating the improved self-driving algorithm development pipeline from fully virtual environments to scaled physical models, to full self-driving cars, potentially leveraging the global developer community for development.

V0128

Time: 18:15-18:30

A New Centrality Measure for Identifying Influential Nodes In Social Networks

Ms. Delel Rhouma and Lotfi Ben Romdhane
MARS Laboratory, ISITCOM, University of sousse, Tunisia

The identification of central nodes has been a key problem in the field of social network analysis. In fact, it is a measure that accounts the popularity or the visibility of an actor within a network. In order to capture this concept, various measures, either sample or more elaborate, has been developed. Nevertheless, many of “traditional” measures are not designed to be applicable to huge data. This paper sets out a new node centrality index suitable for large social network. It uses the amount of vertex’ neighbors and connections between them to characterize a “pivot” node in the graph. We presented experimental results on real data sets which show the efficiency of our proposal.

Poster Session

Venue: Fleming's ROOM 9+10

Time: 13:30pm-18:00pm

*The certificate of poster presentation will be get when collect conference materials at registration desk, please ask chair to sign it for you.

V0015

Graphic Matching based on Shape Contexts and Reweighted Random Walks

Mr. Mingxuan Zhang, Dongmei Niu, Xiuyang Zhao, Mingjun Liu
University of Jinan, Jinan China

Graphic matching is a very critical issue in all aspects of computer vision. In this paper, a new graphics matching algorithm combining shape contexts and reweighted random walks was proposed. On the basis of the local descriptor, shape contexts, the reweighted random walks algorithm was modified to possess stronger robustness and correctness in the final result. Our main process is to use the descriptor of the shape contexts for the random walk on the iteration, of which purpose is to control the random walk probability matrix. We calculate bias matrix by using descriptors and then in the iteration we use it to enhance random walks' and random jumps' accuracy, finally we get the one-to-one registration result by discretization of the matrix. The algorithm not only preserves the noise robustness of reweighted random walks but also possesses the rotation, translation, scale invariance of shape contexts. Through extensive experiments, based on real images and random synthetic point sets, and comparisons with other algorithms, it is confirmed that this new method can produce excellent results in graphic matching.

V0042

A Detail-Preserved and Luminance-Consistent Multi-Exposure Image Fusion Algorithm

Mr. Guanquan Wang and Yue Zhou
Shanghai Jiao Tong University, China

When irradiance across a scene varies greatly, we can hardly get an image of the scene without over- or under-exposure area, because of the constraints of cameras. Multi-exposure image fusion (MEF) is an effective method to deal with this problem by fusing multi-exposure images of a static scene. A novel MEF method is described in this paper. In the proposed algorithm, coarser-scale luminance consistency is preserved by contribution adjustment using the luminance information between blocks; detail-preserved smoothing filter can stitch blocks smoothly without losing details. Experiment results show that the proposed method performs well in preserving luminance consistency and details.

V0047

Compressed Normalized Block Difference for Object Tracking

Mr. Yun Gao, Dengzhuo Zhang, Donglan Cai, Hao Zhou and Ge Lan
Yunnan University, China

Feature extraction is very important for robust and real-time tracking. Compressive sensing provided a technical support for real-time feature extraction. However, all existing compressive tracking were based on compressed Haar-like feature, and how to compress many more excellent high-dimensional features is worth researching. In this paper, a novel compressed normalized block difference feature (CNBD) was proposed.

For resisting noise effectively in a high-dimensional normalized pixel difference feature (NPD), a normalized block difference feature extends two pixels in the original formula of NPD to two blocks. A CNBD feature can be obtained by compressing a normalized block difference feature based on compressive sensing theory, with the sparse random Gaussian matrix as the measurement matrix. The comparative experiments of 7 trackers on 20 challenging sequences showed that the tracker based on CNBD feature can perform better than other trackers, especially than FCT tracker based on compressed Haar-like feature, in terms of AUC, SR and Precision.

V0009

An Adhered-Particle Analysis System Based on Concave Points

Dr. Wencheng Wang, Fengnian Guan, Lin Feng
Weifang University

Particles adhered together will influence the image analysis in computer vision system. In this paper, a method based on concave point is designed. First, corner detection algorithm is adopted to obtain a rough estimation of potential concave points after image segmentation. Then, it computes the area ratio of the candidates to accurately localize the final separation points. Finally, it uses the separation points of each particle and the neighboring pixels to estimate the original particles before adhesion and provides estimated profile images. The experimental results have shown that this approach can provide good results that match the human visual cognitive mechanism.

V0045

Rough Case-Based Reasoning System for Continues Casting

Wenbin Su and **Dr. Zhufeng Lei**
Xi'an Jiaotong University,Xi'an, Shaanxi, China

The continuous casting occupies a pivotal position in the iron and steel industry. The rough set theory and the CBR (case based reasoning, CBR) were combined in the research and implementation for the quality assurance of continuous casting billet to improve the efficiency and accuracy in determining the processing parameters. According to the continuous casting case, the object-oriented method was applied to express the continuous casting cases. The weights of the attributes were calculated by the algorithm which was based on the rough set theory and the retrieval mechanism for the continuous casting cases was designed. Some cases were adopted to test the retrieval mechanism, by analyzing the results, the law of the influence of the retrieval attributes on determining the processing parameters was revealed. A comprehensive evaluation model was established by using the attribute recognition theory. According to the features of the defects, different methods were adopted to describe the quality condition of the continuous casting billet. By using the system, the knowledge was not only inherited but also applied to adjust the processing parameters through the case based reasoning method as to assure the quality of the continuous casting and improve the intelligent level of the continuous casting.

V0061

Age and Gender Estimation using Region-SIFT and Multi-layered SVM

Mr. Hyunduk Kim, Sang-Heon, Lee, Myoung-Kyu Sohn, Byunghun Hwang
DGIST, Republic of Korea

In this paper, we propose an age and gender estimation framework using the region-SIFT

feature and multi-layered SVM classifier. The suggested framework entails three processes. The first step is landmark based face alignment. The second step is the feature extraction step. In this step, we introduce the region-SIFT feature extraction method based on facial landmarks. First, we define sub-regions of the face. We then extract SIFT features from each sub-region. In order to reduce the dimensions of features we employ a Principal Component Analysis (PCA) and a Linear Discriminant Analysis (LDA). Finally, we classify age and gender using a multi-layered Support Vector Machines (SVM) for efficient classification. Rather than performing gender estimation and age estimation independently, the use of the multi-layered SVM can improve the classification rate by constructing a classifier that estimate the age according to gender. Moreover, we collect a dataset of face images, called by DGIST_C, from the internet. A performance evaluation of proposed method was performed with the FERET database, CACD database, and DGIST_C database. The experimental results demonstrate that the proposed approach classifies age and performs gender estimation very efficiently and accurately.

V0014

2.5D Body Estimation via Refined Forest with Field-based Objective

Dr. Jaehwan Kim, HoWon Kim
ETRI , South of Korea

In this paper, we present a 2.5D body region classification method based on the global refinement of random forest. The refinement of random forest provides the reduction of the size of training model while preserving prediction accuracy. We also incorporate the field-inspired objective to the random forest in consideration of the pairwise spatial relationships between neighboring data points. Numerical and visual experiments with artificial 3D data confirm the usefulness of the proposed method.

V0068

A Deep Learning Pipeline for Indian Dance Style Classification

Swati Dewan, **Mr. Shubham Agarwal** and Navjyoti Singh
IIIT Hyderabad, India

In this paper, we address the problem of dance style classification to classify Indian dance or any dance in general. We propose a 3-step deep learning pipeline. First, we extract 14 essential joint locations of the dancer from each video frame, this helps us to derive any body region location within the frame, we use this in the second step which forms the main part of our pipeline. Here, we divide the dancer into regions of important motion in each video frame. We then extract patches centered at these regions. Main discriminative motion is captured in these patches. We stack the features from all such patches of a frame into a single vector and form our hierarchical dance pose descriptor. Finally, in the third step, we build a high level representation of the dance video using the hierarchical descriptors and train it using a Recurrent Neural Network (RNN) for classification. Our novelty also lies in the way we use multiple representations for a single video. This helps us to: (1) Overcome the RNN limitation of learning small sequences over big sequences such as dance; (2) Extract more data from the available dataset for effective deep learning by training multiple representations. Our contributions in this paper are three-folds: (1) We provide a deep learning pipeline for classification of any form of dance; (2) We prove that a segmented representation of a dance video works well with sequence learning techniques for recognition purposes; (3) We extend and refine the ICD dataset and provide a new dataset for evaluation of dance. Our model performs comparable or better in some cases than the state-of-the-art on action recognition benchmarks.

V0123

Laban Movement Analysis to classify Emotions from Motion

Swati Dewan, **Mr. Shubham Agarwal** and Navjyoti Singh
IIIT Hyderabad, India

In this paper, we present the study of Laban Movement Analysis (LMA) to understand basic human emotions from nonverbal human behaviors. While there are a lot of studies on understanding behavioral patterns based on natural language processing and speech processing applications, understanding emotions or behavior from non-verbal human motion is still a very challenging and unexplored field. LMA provides a rich overview of the scope of movement possibilities. These basic elements can be used for generating movement or for describing movement. They provide an inroad to understanding movement and for developing movement efficiency and expressiveness. Each human being combines these movement factors in his/her own unique way and organizes them to create phrases and relationships which reveal personal, artistic, or cultural style. In this work, we build a motion descriptor based on a deep understanding of Laban theory. The proposed descriptor builds up on previous works and encodes experiential features by using temporal windows. We present a more conceptually elaborate formulation of Laban theory and test it in a relatively new domain of behavioral research with applications in human-machine interaction. The recognition of affective human communication may be used to provide developers with a rich source of information for creating systems that are capable of interacting well with humans. We test our algorithm on UCLIC dataset which consists of body motions of 13 non-professional actors portraying angry, fear, happy and sad emotions. We achieve an accuracy of 87.30% on this dataset.

V0096

FPFH-Based Graph Matching for 3D Point Cloud Registration

Mr. Jiapeng Zhao, Chen Li, Lihua Tian and Jihua Zhu
Xi'an Jiaotong University, China

Correspondence detection is a vital step in point cloud registration and it can help getting a reliable initial alignment. In this paper, we put forward an advanced point feature-based graph matching algorithm to solve the initial alignment problem of rigid 3D point cloud registration with partial overlap. Specifically, Fast Point Feature Histograms are used to determine the initial possible correspondences firstly. Next, a new objective function is provided to make the graph matching more suitable for partially overlapping point cloud. The objective function is optimized by the simulated annealing algorithm for final group of correct correspondences. Finally, we present a novel set partitioning method which can transform the NP-hard optimization problem into a $O(n^3)$-solvable one. Experiments on the Stanford and UWA public data sets indicates that our method can obtain better result in terms of both accuracy and time cost compared with other point cloud registration methods.

V0091

Real-time stop sign detection and distance estimation using a single camera

Mr. Wenpeng Wang, Yuxuan Su, Ming Cheng
Boston University, USA

In modern world, the drastic development of driver assistance system has made driving a lot easier than before. In order to increase the safety onboard, a method was proposed to detect

STOP sign and estimate distance using a single camera. In STOP sign detection, LBP-cascade classifier was applied to identify the sign in the image, and the principle of pinhole imaging was based for distance estimation. Road test was conducted using a detection system built with a CMOS camera and software developed by Python language with OpenCV library. Results shows that that the proposed system reach a detection accuracy of maximum of 97.6% at 10m, a minimum of 95.00% at 20m, and 5% max error in distance estimation. The results indicate that the system is effective and has the potential to be used in both autonomous driving and advanced driver assistance driving systems.

V0085

Researches of Fruit Quality Prediction Model Based on Near Infrared Spectrum

Dr. Yulin Shen, Lian Li
School of Information Science and Engineering, Lanzhou University, Lanzhou, China

With the improvement in standards for food quality and safety, people pay more attention to the internal quality of fruits, therefore the measurement of fruit internal quality is increasingly imperative. In general, nondestructive soluble solid content (SSC) and total acid content (TAC) analysis of fruits is vital and effective for quality measurement in global fresh produce markets, so in this paper, we aim at establishing a novel fruit internal quality prediction model based on SSC and TAC for Near Infrared Spectrum. Firstly, the model of fruit quality prediction based on PCA + BP neural network, PCA + GRNN network, PCA + BP adaboost strong classifier, PCA + ELM and PCA + LS_SVM classifier are designed and implemented respectively; then, in the NSCT domain, the median filter and the Savitzky-Golay filter are used to preprocess the spectral signal, Kennard-Stone algorithm is used to automatically select the training samples and test samples; thirdly, we achieve the optimal models by comparing 15 kinds of prediction model based on the theory of multi-classifier competition mechanism, specifically, the non-parametric estimation is introduced to measure the effectiveness of proposed model, the reliability and variance of nonparametric estimation evaluation of each prediction model to evaluate the prediction result, while the estimated value and confidence interval regard as a reference, the experimental results demonstrate that this model can better achieve the optimal evaluation of the internal quality of fruit; finally, we employ cat swarm optimization to optimize two optimal models above obtained from non-parametric estimation, empirical testing indicates that the proposed method can provide more accurate and effective results than other forecasting methods.

V0024

Comparison of classification algorithms for various methods of preprocessing radar images of the MSTAR base

Mr. Aleksandr. A. Borodinov, Vladislav. V. Myasnikov
Samara National Research University, Samara, Russia

The present work is devoted to comparing the accuracy of the known qualification algorithms in the task of recognizing local objects on radar images for various image preprocessing methods. Preprocessing involves speckle noise filtering and normalization of the object orientation in the image by the method of image moments and by a method based on the Hough transform. In comparison, the following classification algorithms are used: Decision tree; Support vector machine, AdaBoost, Random forest. The principal component analysis is used to reduce the dimension. The research is carried out on the objects from the base of radar images MSTAR. The paper presents the results of the conducted studies.

V0004

Optimization of the hierarchical interpolator for image compression

Assoc. Prof. M.V. Gashnikov
Samara National Research University, Russia

Hierarchical interpolation of images is investigated in the problem of image compression. A new approach is proposed for optimizing the adaptive interpolator for hierarchical compression. This approach is based on optimizing the entropy of the compressed signal. This approach is more adequate to the compression problem than the known approach based on optimization of the interpolation error. The optimization algorithm for the adaptive interpolator is proposed on the basis of the proposed approach. The theoretical estimation of the computational complexity of the proposed interpolator is calculated. A comparison of this complexity with the complexity of other interpolators is performed. The advantage of the proposed interpolator over known interpolators is investigated experimentally. The win is calculated according to the size of the archive file. Recommendations for the use of the proposed interpolator are formulated.

V3001

Formation of the Image on the Receiver of Thermal Radiation

Tatiana A Akimenko
Tula State University 92 , Leninal prospect, Tula, 300012, Russia

The formation of the thermal picture of the observed scene with the verification of the quality of the thermal images obtained is one of the important stages of the technological process that determine the quality of the thermal imaging observation system. In this article, we propose to consider a model for the formation of a thermal picture of a scene, which must take into account: the features of the object of observation as the source of the signal; signal transmission through the physical elements of the thermal imaging system that produce signal processing at the optical, photoelectronic and electronic stages, which determines the final parameters of the signal and its compliance with the requirements for thermal information and measurement systems.

V0087

Distribution majorization of Corner Points by Reinforcement Learning for Moving Object Detection

Mr. Hao Wu, Hao Yu, Dongxiang Zhou and Yongqiang Cheng
National University of Defense Technology, China

Corner points play an important role in moving object detection, especially in the case of free-moving camera. Corner points provide more accurate information than other pixels and reduce the computation which is un- necessary. Previous works only use intensity information to locate the corner points, however, the information that former and the last frames provided also can be used. We utilize the information to focus on more valuable area and ignore the invaluable area. The proposed algorithm is based on reinforcement learning, which regards the detection of corner points as a Markov process. In the Markov model, the video to be detected is regarded as environment, the selections of blocks for one corner point are regarded as actions and the performance of detection is regarded as state. Corner points are assigned to be the blocks which are seperated from original whole image. Experimentally, we select a conventional method which uses marching and Random Sample Con- sensus

algorithm to obtain objects as the main framework and utilize our algorithm to improve the result. The comparison between the conventional method and the same one with our algorithm show that our algorithm reduce 70% of the false detection.

V0075

Enhanced Online Convolutional Neural Networks for Object Tracking

Dengzhuo Zhang, **Mr. Yun Gao**, Hao Zhou, Tianwen Li
Yunnan University, China

In recent several years, object tracking based on convolution neural network has gained more and more attention. The initialization and update of convolution filters can directly affect the precision of object tracking effective. In this paper, a novel object tracking via an enhanced online convolution neural network without offline training is proposed, which initializes the convolution filters by a k-means++ algorithm and updates the filters by an error back-propagation. The comparative experiments of 7 trackers on 15 challenging sequences showed that our tracker can perform better than other trackers in terms of AUC and precision.

V0097

A Comparative Analysis of Image Features between Weave Embroidered Thangka and Piles Embroidered Thangka

Mr. Zhenjiang Li, Weilan Wang
Northwest Minzu University, China

Thangka is a treasure of Tibetan culture. In its digital protection, most of the current research focuses on the content of Thangka images, not the fabrication process. For silk embroidered Thangka of "Guo Tang", there are two craft methods, namely, weave embroidered and piles embroidered. The local texture of weave embroidered Thangka is rough, and that of piles embroidered Thangka is more smooth. In order to distinguish these two kinds of fabrication processes from images, a effectively segmentation algorithm of color blocks is designed firstly, and the obtained color blocks contain the local texture patterns of Thangka image; Secondly, the local texture features of the color block are extracted and screened; Finally, the selected features are analyzed experimentally. The experimental analysis shows that the proposed features can well reflect the difference between methods of weave embroidered and piles embroidered.

V0023

A review on three main paradigms of Simultaneous Localization and Mapping (SLAM) problem

Mr. Vandad Imani, Keijo Haataja and Pekka Toivanen
UNIVERSITY OF EASTERN FINLAND, FINLAND

Simultaneous Localization and Mapping (SLAM) is one of the most challenging research areas within computer and machine vision for automated scene commentary and explanation. The SLAM technique has been a developing research area in the robotics context during recent years. By utilizing the SLAM method robot can estimate the different positions of the robot at the distinct points of time which can indicate the trajectory of robot as well as generate a map of the environment. SLAM has unique traits which are estimating the location of robot and building a map in the various types of environment. SLAM is effective in different types of environment such as indoor, outdoor district, Air, Underwater, Underground and Space. Several approaches have been investigated to use SLAM technique in distinct environments. The purpose of this paper is to provide an accurate perceptive review of case

history of SLAM relied on laser/ultrasonic sensors and camera as perception input data. In addition, we mainly focus on three paradigms of SLAM problem with all its pros and cons. In the future, use intelligent methods and some new idea will be used on visual SLAM to estimate the motion intelligent underwater robot and building a feature map of marine environment.

V0039

Hand motion modeling for psychology analysis in job interview using Optical Flow-History Motion Image (OF-HMI)

Ms. Mariem Gnouma, Intissar KHALIFA, Ridha EJBALI, Mourad ZAIED
Research Team in Intelligent Machines, National Engineering School of Gabes, Tunisia

To survive the competition, companies always think about having the best employees. The selection is depended on the answers to the questions of the interviewer and the behavior of the candidate during the interview session. The study of this behavior is always based on a psychological analysis of the movements accompanying the answers and discussions. Few techniques are proposed until today to analyze automatically candidate's non verbal behavior. This paper is a part of a work psychology recognition system; it concentrates in spontaneous hand gesture which is very significant in interviews according to psychologists. We propose motion history representation of hand based on an hybrid approach that merges optical flow and history motion images. The optical flow technique is used firstly to detect hand motions in each frame of a video sequence. Secondly, we use the history motion images (HMI) to accumulate the output of the optical flow in order to have finally a good representation of the hand's local movement in a global temporal template.

V0038

An Optimal Approach for Deployment Sensors in WSN

Ms. Mariem Gnouma, Hamida Grari, Faten Hajjej, Ridha Ejbali, and Mourad Zaied
Research Team in Intelligent Machines (RTIM) Gabes, Tunisia

Wireless Sensor Networks (WSNs) is considered the basis of all architectures of IoT applications. Those networks are deployed according to the needs of each domain. The goals of the proposed approaches of deployment are the lifetime maximization of each device. The major challenge is to find a tradeoff between desired requirements for the lifetime, coverage and cost with limited available resources. In this paper, we present a new method for the deployment of WSN that tries to increase the lifetime of the network. A set of optimum position are presented after their calculation using a GPS. The objective is to locate our sensors in these optimum positions. We try to minimize the energy consumption by decreasing the distance travelled by the sensors to reach their optimum positions.

V0110

Faulty Node Detection in Wireless Sensor Networks using a Recurrent Neural Network

Ms. Mariem Gnouma, Jamila Atiga, Nour El Houda Mbarki, Ridha Ejbali1, and Mourad Zaied
University of Gabes, Faculty of Sciences of Gabes (FSG)

The wireless sensor networks (WSN) consist of a set of sensors that are more and more used in surveillance applications on a large scale in different areas: military, Environment, Health ... etc. Despite the minimization and the reduction of the manufacturing costs of the sensors, they can operate in places difficult to access without the possibility of reloading of

battery, they generally have limited resources in terms of power of emission, of processing capacity, data storage and energy. These sensors can be used in a hostile environment, such as, for example, on a field of battle, in the presence of fires, floods, earthquakes. In these environments the sensors can fail, even in a normal operation. It is therefore necessary to develop algorithms tolerant and detection of defects of the nodes for the network of sensor without wires, therefore, the faults of the sensor can reduce the quality of the surveillance if they are not detected. The values that are measured by the sensors are used to estimate the state of the monitored area. We used the Non-linear Auto-Regressive with eXogeneous (NARX), the recursive architecture of the neural network, to predict the state of a node of a sensor from the previous values described by the functions of time series. The experimental results have verified that the prediction of the State is enhanced by our proposed model.

V0008

A Low Cost Machine Vision System for the Recognition and Sorting of Small Parts

G. Barea, **Prof. B.W. Surgenor** and V.D. Chauhan
Queen's University, Kingston, Canada

An automated machine vision-based system for the recognition and sorting of small parts was designed, assembled and tested. The system was developed to address a need to expose engineering students to the issues of machine vision and assembly automation technology, with readily available and relatively low cost hardware and software. This paper outlines the design of the system and presents experimental performance results. Three different styles of plastic gears, together with three different styles of defective gears, were used to test the system. A pattern matching tool was used for part classification. Nine experiments were conducted to demonstrate the effects of changing various hardware and software parameters, including: conveyor speed, gear feed rate, classification, and identification score thresholds. If was found that the system could achieve a maximum system accuracy of 95% at a feed rate of 60 parts/min, for a given set of parameter settings. Future work will be looking at the effect of lighting.

V0027

Compressed Multi-Block Local Binary Pattern for Object Tracking

Tianwen LI, **Mr. Yun GAO**, Lei ZHAO, Hao ZHOU
Yunnan University, China

Both robustness and real-time are very important for the application of object tracking under a real environment. The focused trackers based on deep learning are difficult to satisfy with the real-time of tracking. Compressive sensing provided a technical support for real-time tracking. In this paper, an object can be tracked via a multi-block local binary pattern feature. The feature vector was extracted based on the multi-block local binary pattern feature, which was compressed via a sparse random Gaussian matrix as the measurement matrix. The experiments showed that the proposed tracker ran in real-time and outperformed the existed compressive trackers based on Haar-like feature on many challenging video sequences in terms of accuracy and robustness.

V0124A

Machine vision system for three-dimensional imaging in adjustable narrow spectral bands

Assoc. Prof. Alexander Machikhin, Vladislav Batshev, Alexander Naumov, Demid Khokhlov, Vitold Pozhar and Alexey Gorevoy
Russia Scientific and Technological Center of Unique Instrumentation of Russian Academy of Sciences

We propose a new machine vision system for three-dimensional (3D) imaging in arbitrary spectral intervals. It is based on a simultaneous diffraction of two convergent stereoscopic light beams on acoustic waves in uniaxial birefringent crystals. In this study, we discuss the influence acousto-optic interaction configuration on the image quality and the peculiarities of geometrical and spectral calibration. Experimental results demonstrate the effectiveness of the proposed approach for 3D spectral imaging and precise 3D measurements. Proposed system may be implemented as an ultra-compact embedded component and is perspective for many imaging applications. Among the most promising applications are machine vision and microscopy, where the spectral filtration in the particular spectral intervals produces contrast patterns of physical, chemical and other properties of the inspected objects. The multi- and hyperspectral imaging reveals important additional features in comparison with conventional color visualization.

Listeners

*The certificate of listener will be available when collect conference materials at registration desk, please ask chair to sign it for you.

Lis 1

Dalila Salamani
University of Geneva, Switzerland

Lis 2

Dzmitry Piliutsik
Belarus, Minsk

Lis 3

Katsiaryna Matusevich
Belarus, Minsk

Lis 4

Raman Tsialiatnikau
Belarus, Minsk

Lis 5

Vladimir Arlazarov
Russia,117574,Moscow, Golubinskaya 7-2-349

Lis 6

Paolo Miocchi
MARBILab c/o Fondazione Santa Lucia Via Ardeatina, 306
00179 Roma - Italy

Lis 7

Johan DEBAYLE
Ecole Nationale Supérieure des Mines, Saint-Etienne, FRANCE

Author Index

Name	ID	Page	Session
A			
A. Cedola	V0111	41	S6
A. Lukoyanov	V0121	29	S3
Ahmet ARSLAN	V0002	26	S2
Aleksander Lynchenko	V0090	33	S4
Aleksandr Yurievich Shkanaev	V0006	19	S1
Aleksandr Yurievich Shkanaev	V0066	37	S5
Aleksandrs Levinskis	V0135	38	S5
Aleksei Vladimirovich Panchenko	V0066	37	S5
Aleksei Vladimirovich Panchenko	V0006	19	S1
Alessandro Colombo	V0013	23	S2
Alexander A. Tereshin	V0089	32	S4
Alexander Kozharinov	V0119	31	S4
Alexander Privalov	V0073	49	S8
Alexander Sheshkus	V0138	25	S2
Alexander V. Gayer	V0105	50	S8
Alexander V. Sheshkus	V0105	50	S8
Alexander Zhukovsky	V0101	47	S7
Alexey Buzmakov	V0113	36	S5
Alexey Buzmakov	V0137	48	S8
Alexey V. Savchik	V0044	29	S3
Alima Damak	V0077	27	S3
Aliy Delov	V0132	33	S4
Anastasia Ingacheva	V0138	25	S2
Anastasia Ingacheva	V0137	48	S8
Anastasia Ingacheva	V0011	35	S5
Anastasia Ingacheva	V0107	40	S6
Andreas Pichler	V0099A	42	S6
Andrej Satnik	V0031	44	S7
Andrey Gladkov	V0065	38	S5
Andrey Gladkov	V0011	35	S5
Anne van Rossum	V0048	48	S8
Anton Kummert	V0013	23	S2
Arlazarov V.L.	V0112	22	S1
Arvind C S	V0108	24	S2
Aya Kubota	V0026	44	S7
Azeddine ELHASSOUNY	V0088	28	S3
B			
Baomin Zhang	V0017	37	S5
Basel Solaiman	V0084	28	S3
Betül UZBAŞ	V0002	26	S2
Bezmaternykh P.V.	V0112	22	S1
Boris Savelyev	V0098	30	S3
C			
Chadia Khraief	V0130	24	S2
Chi-Tai Vong	V0060	27	S3
Chokri Ben Amar	V0050	21	S1
Christian Fuchs	V0011	35	S5
Christoph Heindl	V0102	42	S6
Claudio Ferrero	V0101	47	S7
Cong Li	V0066	37	S5
Cunyue Lu	V0066	37	S5
D			
D. Nikolaev	V0121	29	S3
D. Sidorchuk	V0074	46	S7
Dalila Salamani	V0049	45	S7
Damian Borys	V0115	46	S7
Daniil Tropin	V0098	30	S3
Delel Rhouma	V0128	51	S8
Denis Zolotov	V0113	36	S5
Didier Stricker	V0107	40	S6
Dietrich Paulus	V0011	35	S5
Dmitriy Novikov	V0104	46	S7
Dmitry Abulkhanov	V0079	30	S3
Dmitry Ilin	V0104	46	S7
Dmitry Nikolaev	V0101	47	S7
Dmitry Nikolaev	V0092	31	S4
Dmitry Nikolaev	V0011	35	S5
Dmitry Nikolaev	V0072	40	S6
Dmitry Nikolaev	V0074	46	S7
Dmitry Nikolaev	V0101	47	S7
Dmitry Nikolaev	V0079	30	S3
Dmitry Nikolaev	V0138	25	S2
Dmitry Nikolaev	V0017	37	S5
Dmitry Nikolaev Alexey Savchik	V0079	30	S3
Dmitry Nikolaevich Putintsev	V0051	36	S5
Dmitry P. Nikolaev	V0044	29	S3
Dmitry Polevoy	V0104	46	S7
Dmitry Sidorchuk	V0054	29	S3
Dmitry Valerevich Polevoy	V0006	19	S1
Dorra Sellami	V0058	49	S8

Dorra Sellami	V0060	27	S3
E			
Ebroul Izquierdo	V0031	44	S7
Ehsan Kiani	V0139	36	S5
Ehsanollah Kabir	V0056	40	S6
Elena Andreeva	V0093	22	S1
Elena Limonova	V0098	30	S3
Eugene Larkin	V0073	49	S8
Evgenii Semenishchev	V0134	34	S4
Evgeny Semenishchev	V0133	47	S7
Evgeny Shvets	V0054	29	S3
Evgeny Shvets	V0116	25	S2
Eya Mezghani	V0052	21	S1
F			
Fahad Mahmood	V0054	29	S3
Fahad Mahmood	V0082	41	S6
Fahad Mehmood	V0082	41	S6
Faouzi Benzarti	V0130	24	S2
Farzin Ghorban	V0057	23	S2
Fatma Zehra Göğüş	V0029	19	S1
Fernando Rubio Perona	V0051	36	S5
Feten Besbes	V0077	27	S3
Florence Cloppet	V0058	49	S8
Frank Neuhaus	V0011	35	S5
G			
G. Campi	V0099A	42	S6
Gang Zhang	V0053	49	S8
Geraldin NANFACK	V0088	28	S3
Gernot Stubl	V0102	42	S6
Giovanni Savino	V0063	32	S4
Giovanni Savino	V0114	33	S4
Gongguo Tang	V0035	20	S1
Gülay Tezel	V0029	19	S1
Gustavo Gil	V0063	32	S4
Gustavo Gil	V0114	33	S4
H			
H.Jaap van den Herik	V0137	48	S8
Hai Xiang Lin	V0048	48	S8
Hamid Amiri	V0130	24	S2
Hamit Altıparmak	V0139	36	S5
Hamrouni Kamel	V0025	44	S7
Heiko Sparenberg	V0035	20	S1
Henri Nicolas	V0052	21	S1
Henryk Palus	V0104	46	S7
Hongyu Wang	V0017	37	S5
Hossein aharvand	V0105	50	S8
I			
I. Bukreeva	V0099A	42	S6
I. Konovalenko	V0121	29	S3
Igor Faradjev	V0082	41	S6
Igor Janiszewski	V0098	30	S3
Igor Mamay	V0079	30	S3
Igor Polyakov	V0116	25	S2
Ing. Deitrich Paulus	V0050	21	S1
Ingars Ribners	V0125	51	S8
Irina Kunina	V0065	38	S5
Ivan Andeevich Konovalenko	V0051	36	S5
Ivan Konovalenko	V0079	30	S3
Ivan Yakimchuk	V0011	35	S5
J			
Javaid Iqbal	V0054	29	S3
Javaid Iqbal	V0056	40	S6
Javier Marín	V0057	23	S2
Jawad Mahmood	V0054	29	S3
Jiaquan Li	V0053	49	S8
Jignesh S. Bhatt	V0070	24	S2
Jlassi Hajer	V0025	44	S7
Joao Batista Florindo	V0043	44	S7
Johan Debayle	V0111	41	S6
Johan Dubbeldam	V0048	48	S8
Jose Miguel Puerta Callejón	V0051	36	S5
Josef Scharinger	V0099A	42	S6
Julia Shemiakina	V0103	41	S6
Julia Shemiakina	V0101	47	S7
Junchao Tu	V0080	50	S8
K			
K.D. Joshi Prof. B.W. Surgenor	V0007	35	S5
Kamel Hamrouni	V0084	28	S3
Kamil Dimililer	V0113	36	S5
Kamil Gorczewski	V0115	46	S7
Kenta Morishita	V0072	40	S6
Konstantin Bulatov	V0092	31	S4
Konstantin Bulatov	V0090	33	S4
Krisjanis Nesenbergs	V0125	51	S8
Krokhina Darya Alekseevna	V0128	51	S8
Kumar Vishal	V0108	24	S2
L			
L. Massimi	V0111	41	S6
László Lefkovits	V0069	21	S1
Laurent Wendling	V0053	49	S8
Lev Teplyakov	V0065	38	S5
Liyan Zhang	V0080	50	S8

Lotfi Ben Romdhane	V0128	51	S8
M			
M. Fratini	V0099A	42	S6
M. Julia Flores Gallego	V0051	36	S5
Maha Charfeddine	V0052	21	S1
Mahboobeh Jannesari	V0106	50	S8
Manjunath V. Joshi	V0070	24	S2
Manjunath V. Joshi	V0122	51	S8
Manu Sharma	V0070	24	S2
Marco Pierini	V0063	32	S4
Marco Pierini	V0090	33	S4
Marek Kijonka	V0104	46	S7
Mariem Gnouma	V0109	25	S2
Marina Chukalina	V0107	40	S6
Marina Chukalina	V0113	36	S5
Marina Chukalina	V0101	47	S7
Mariusz Frackiewicz	V0104	46	S7
Marwa Hmida	V0084	28	S3
Marwa Khatrouch	V0109	25	S2
Masatoshi Ishikawa	V0135	38	S5
Maxim Minkin	V0132	33	S4
Mehdi Habibzadeh	V0106	50	S8
Mehdi Totonchi	V0105	50	S8
Miguel Ángel Martínez-Del-Amor	V0046	20	S1
Minh-Triet Tran	V0002	26	S2
Mohamad Al Shahadat	V0139	36	S5
Mohammad Reza Soheili	V0056	40	S6
Mohammad Reza Yousefi	V0056	40	S6
Morosanu Bogdan Andrei Lemnaru Camelia	V0034	34	S4
Mourad Zaied	V0116	25	S2
N			
Natalya Kotova	V0058	49	S8
Natalya Skoryukina	V0103	41	S6
Nauris Dorbe	V0125	51	S8
Naveed Ejaz	V0046	20	S1
Nikita Razumnuy	V0119	31	S4
Nikolaev D.P.	V0112	22	S1
Nima Hatami	V0111	41	S6
Norhene Gargouri	V0077	27	S3
Nouha Dammak	V0132	33	S4
O			
Oleg Slavin	V0092	31	S4
Oleg Slavin	V0069	21	S1
Ouazaa Hibet-Allah	V0025	44	S7
R			
Rachid OULD HAJ THAMI	V0088	28	S3
Rawya Al-Akam	V0050	21	S1
Richard Orjesek	V0031	44	S7
Ridha Ejbali	V0109	25	S2
Rinat Nailevish Sadekov	V0051	36	S5
Ritesh Mishra	V0108	24	S2
Roberts Kadikis	V0012	32	S4
S			
Sabrina Amrouche	V0049	45	S7
Sana Boussetta	V0077	27	S3
Sebastian Budzan	V0013	23	S2
Seiichi Gohshi	V0026	44	S7
Sei-ichiro Kamata	V0040	42	S6
Sergey A. Usilin	V0089	32	S4
Sikan Tang	V0053	49	S8
Simina Emerich	V0069	21	S1
Simone Piantini	V0114	33	S4
Soumaya Trabelsi Ben Ameur	V0058	49	S8
Szidónia Lefkovits	V0069	21	S1
T			
Temudzhin Manzhikov	V0093	22	S1
Temudzhin Manzhikov	V0098	30	S3
Thomas Pönitz	V0102	42	S6
Timofey Chernov	V0119	31	S4
Timur Khanipov	V0107	40	S6
Timur Khanipov	V0017	37	S5
Tomohiko Hayakawa	V0072	40	S6
U			
Umair Ali Khan	V0046	20	S1
Umair Munir	V0082	41	S6
Uryi Igorevish Minkin	V0066	37	S5
V			
V. Volkov	V0074	46	S7
V.D. Chauhan	V0007	35	S5
Vaibhav Amit Patel	V0122	51	S8
Valeriy Krivtsov	V0090	33	S4
Venugopal Gundimeda	V0070	24	S2
Viacheslav Voronin	V0134	34	S4
Victor Asadchikov	V0113	36	S5
Victor Asadchikov	V0101	47	S7
Victor Prun	V0137	48	S8
Victoria A. Sablina	V0044	29	S3
Vinh Ton-That	V0060	27	S3
Vladimir Arlazarov	V0103	41	S6
Vladimir Arlazarov	V0092	31	S4
Vladimir Dub	V0133	47	S7

Vladimir Frantc	V0134	34	S4	Xun Zhao	V0017	37	S5
Vladimir V. Arlazarov	V0093	22	S1	Y			
Vladimir V. Arlazarov	V0089	32	S4	Yann Gavet	V0111	41	S6
Vladimir V. Arlazarov	V0092	31	S4	Yassine BenAyed	V0132	33	S4
Vladislav Kotov	V0073	49	S8	Yitian Wang	V0040	42	S6
Vyacheslav Voronin	V0133	47	S7	Youye Xie	V0035	20	S1
W				Yu Su	V0057	23	S2
William Hoff	V0035	20	S1	Yulia S. Chernyshova	V0105	50	S8
Wojciech Serafin	V0115	46	S7	Yushi Moko	V0072	40	S6
X				Z			
Xuan-Truong Nguyen-Dao	V0060	27	S3	Zahra Rezaei	V0106	50	S8

ONE DAY VISIT

Vienna is the capital and largest city of Austria and one of the nine states of Austria. Apart from being regarded as the City of Music because of its musical legacy, Vienna is also said to be "The City of Dreams" because it was home to the world's first psycho-analyst – Sigmund Freud. The city's roots lie in early Celtic and Roman settlements that transformed into a Medieval and Baroque city, and then the capital of the Austro-Hungarian Empire. It is well known for having played an essential role as a leading European music centre, from the great age of Viennese Classicism through the early part of the 20th century. The historic centre of Vienna is rich in architectural ensembles, including Baroque castles and gardens, and the late-19th-century Ringroad lined with grand buildings, monuments and parks.

Schloss Schonbrunn (美泉宫)

Schloss Schönbrunn is a Baroque art building in the southwest of the Austrian capital. It was the Holy Roman Empire, the Austrian Empire, the Austro-Hungarian Empire and the Habsburg dynasty family Is now the most prestigious tourist attraction in Vienna.

Schönbrunn's name comes from the Holy Roman Empire Matthias (1612 to 1619 reign), legended in 1612 he was hunting to Meidling and Hietzing between the Fort Kate, drink here spring, fresh Ganlie , Then named this spring as 'spring', then 'Meiquan' to become the name of the region. In 1743, the Queen Mary Maria Treesia ordered the construction of the magnificent Schönbrunn and Baroque gardens, with a total area of 26,000 square meters, second only to the French Versailles.

Vienna State Opera （维也纳歌剧院）

The Vienna State Opera (German: Wiener Staatsoper) is an Austria opera house and opera company based in Vienna, Austria. It was originally called the Vienna Court Opera (Wiener Hofoper). In 1920, with the replacement of the Habsburg Monarchy by the First Austrian Republic, it was renamed the Vienna State Opera. The members of the Vienna Philharmonicare recruited from its orchestra.

St. Stephen's Cathedral （圣斯蒂芬大教堂）

St. Stephen's Cathedral (more commonly known by its German title: Stephansdom) is the mother church of the Roman Catholic Archdiocese of Vienna and the seat of the Archbishop of Vienna, Christoph Cardinal Schönborn, OP. The current Romanesque and Gothic form of the cathedral, seen today in the Stephansplatz, was largely initiated by Duke Rudolf IV(1339–1365) and stands on the ruins of two earlier churches, the first a parish church consecrated in 1147. The most important religious building in Vienna, St. Stephen's Cathedral has borne witness to many important events in Habsburg and Austrian history and has, with its multi-coloured tile roof, become one of the city's most recognizable symbols.

Hofburg (霍夫堡宫)

The **Hofburg** is the former imperial palace in the centre of Vienna, Austria. Built in the 13th century and expanded in the centuries since, the palace has been the seat of power of the Habsburg dynasty rulers, and today the official residence and workplace of the President of Austria. It was the principal imperial winter residence, as Schönbrunn Palace was the summer residence.

Kunsthistorisches Museum （艺术史博物馆）

The **Kunsthistorisches Museum** (English: "Museum of Art History", also often referred to as the "Museum of Fine Arts") is an art museum in Vienna, Austria. Housed in its festive palatial building on Ringstraße, it is crowned with an octagonal dome. The term *Kunsthistorisches Museum* applies to both the institution and the main building. It is the largest art museum in the country.

MEMO