

Adaptive spatial resolution selection for stereoscopic video compression with MV-HEVC: A frequency based approach

One approach for stereoscopic video compression is to down sample the content prior to encoding and up sample it to the original spatial resolution after decoding. In this study it is shown that the ratio by which the content should be rescaled is sequence dependent. Hence, a frequency based method is introduced enabling fast and accurate estimation of the best down sampling ratio for different stereoscopic video clips. It is shown that exploiting this approach can bring 3.38% delta bitrate reduction over five camera-captured sequences.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: Video, Research Community on Data-to-Decision (D2D), Nokia

Contributors: Aflaki, P., Hannuksela, M. M., Gabbouj, M.

Number of pages: 4

Pages: 267-270

Publication date: 5 Feb 2015

Host publication information

Title of host publication: 2014 IEEE International Symposium on Multimedia, ISM 2014, 10-12 Dec. 2014, Taichung

Publisher: The Institute of Electrical and Electronics Engineers, Inc.

ISBN (Print): 9781479943111

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Human-Computer Interaction, Software, Media Technology

Keywords: frequency power spectrum, MVC, objective quality metrics, resolution adjustment

DOIs:

10.1109/ISM.2014.11

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

An evaluation framework for cross-platform mobile app development tools: A case analysis of adobe PhoneGap framework

The 'App economy' is a highly lucrative and competitive market for independent software vendors as it potentially offers an easy highway to reach millions of users. However, the mobile application landscape is scattered and an application developer has to publish the software for several different platforms to be able to serve a majority of smartphone users. Therefore, a bunch of cross-development tools have been offered to simplify this workload. In this paper, we present an evaluation framework for comparing different cross-development tools. We use this framework to evaluate Adobe PhoneGap tool against native development in Android and Windows Phone platforms. The results of a case study reveal that while the cross-platform technique was easy to use, the appearance and usability of the app was mediocre at its best. The business impacts of these are also discussed.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Turun Yliopisto/Turun Biomateriaalikeskus

Contributors: Ahti, V., Hyrynsalmi, S., Nevalainen, O.

Number of pages: 8

Pages: 41-48

Publication date: 23 Jun 2016

Host publication information

Title of host publication: Computer Systems and Technologies 17th International Conference, CompSysTech 2016 - Proceedings

Volume: 1164

Publisher: Association for Computing Machinery

ISBN (Electronic): 9781450341820

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Cross-platform development, Hybrid mobile app, Mobile application, Multi-platform

DOIs:

10.1145/2983468.2983484

URLs:

<http://www.scopus.com/inward/record.url?scp=85001085934&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 85001085934

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Keyframe-based video summarization with human in the loop

In this work, we focus on the popular keyframe-based approach for video summarization. Keyframes represent important and diverse content of an input video and a summary is generated by temporally expanding the keyframes to key shots which are merged to a continuous dynamic video summary. In our approach, keyframes are selected from scenes that represent semantically similar content. For scene detection, we propose a simple yet effective dynamic extension of a video Bag-of-Words (BoW) method which provides over segmentation (high recall) for keyframe selection. For keyframe selection, we investigate two effective approaches: local region descriptors (visual content) and optical flow descriptors (motion content). We provide several interesting findings. 1) While scenes (visually similar content) can be effectively detected by region descriptors, optical flow (motion changes) provides better keyframes. 2) However, the suitable parameters of the motion descriptor based keyframe selection vary from one video to another and average performances remain low. To avoid more complex processing, we introduce a human-in-the-loop step where user selects keyframes produced by the three best methods. 3) Our human assisted and learning-free method achieves superior accuracy to learning-based methods and for many videos is on par with average human accuracy.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Research group: Vision

Contributors: Ainasoja, A. E., Hietanen, A., Lankinen, J., Kämäräinen, J.

Number of pages: 10

Pages: 287-296

Publication date: 2018

Host publication information

Title of host publication: VISIGRAPP 2018 - Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications

Volume: 4

Publisher: SCITEPRESS

ISBN (Electronic): 9789897582905

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Graphics and Computer-Aided Design, Artificial Intelligence

Keywords: Optical flow descriptors, Region descriptors, Video summarization, Visual bag-of-words

DOIs:

10.5220/0006619202870296

Bibliographical note

INT=sgn,"Lankinen, Jukka"

Source: Scopus

Source ID: 85047872595

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Smartphone teleoperation for self-balancing telepresence robots

Self-balancing mobile platforms have recently been adopted in many applications thanks to their light-weight and slim build. However, inherent instability in their behaviour makes both manual and autonomous operation more challenging as compared to traditional self-standing platforms. In this work, we experimentally evaluate three teleoperation user interface approaches to remotely control a self-balancing telepresence platform: 1) touchscreen button user interface, 2) tilt user interface and 3) hybrid touchscreen-tilt user interface. We provide evaluation in quantitative terms based on user trajectories and recorded control data, and qualitative findings from user surveys. Both quantitative and qualitative results support our finding that the hybrid user interface (a speed slider with tilt turn) is a suitable approach for smartphone-based teleoperation of self-balancing telepresence robots. We also introduce a client-server based multi-user telepresence architecture using open source tools.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences

Contributors: Ainasoja, A. E., Pertuz, S., Kämäräinen, J.

Number of pages: 8

Pages: 561-568

Publication date: 2019

Host publication information

Title of host publication: VISIGRAPP 2019 - Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications

Publisher: SCITEPRESS

Editors: Kerren, A., Hurter, C., Braz, J.

ISBN (Electronic): 9789897583544

ASJC Scopus subject areas: Computer Science Applications, Computer Vision and Pattern Recognition, Computer Graphics and Computer-Aided Design

Keywords: Teleoperation, Telepresence, User Interface

Electronic versions:

VISAPP_2019_199

DOIs:

10.5220/0007406405610568

URLs:

<http://urn.fi/URN:NBN:fi:tty-201908282038>

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

TraQuMe: A tool for measuring the gaze tracking quality

Consistent measuring and reporting of gaze data quality is important in research that involves eye trackers. We have developed TraQuMe: a generic system to evaluate the gaze data quality. The quality measurement is fast and the interpretation of the results is aided by graphical output. Numeric data is saved for reporting of aggregate metrics for the whole experiment. We tested TraQuMe in the context of a novel hidden calibration procedure that we developed to aid in experiments where participants should not know that their gaze is being tracked. The quality of tracking data after the hidden calibration procedure was very close to that obtained with the Tobii's T60 trackers built-in 2 point, 5 point and 9 point calibrations.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA), School of Management (JKK)

Contributors: Akkil, D., Isokoski, P., Kangas, J., Rantala, J., Raisamo, R.

Number of pages: 4

Pages: 327-330

Publication date: 2014

Host publication information

Title of host publication: Proceedings of the Symposium on Eye Tracking Research and Applications, ETRA 2014

Publisher: Association for Computing Machinery

ISBN (Print): 9781450327510

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems

Keywords: Gaze interaction, Gaze tracking

DOIs:

10.1145/2578153.2578192

URLs:

<http://www.scopus.com/inward/record.url?scp=84899688722&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84899688722

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Viewing simulation of integral imaging display based on wave optics

We present an accurate model of integral imaging display based on wave optics. The model enables accurate characterization of the display through simulated perceived images by the human visual system. Thus, it is useful to investigate the capabilities of the display in terms of various quality factors such as depth of field and resolution, as well as delivering visual cues such as focus. Furthermore, due to the adopted wave optics formalism, simulation and analysis of more advanced techniques such as wavefront coding for increased depth of field are also possible.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing

Contributors: Akpinar, U., Sahin, E., Gotchev, A.

Publication date: 1 Oct 2018

Host publication information

Title of host publication: 2018 - 3DTV-Conference : The True Vision - Capture, Transmission and Display of 3D Video, 3DTV-CON 2018

Publisher: IEEE

Article number: 8478568

ISBN (Electronic): 9781538661253

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Networks and Communications, Computer Vision and Pattern Recognition, Human-Computer Interaction, Electrical and Electronic Engineering

Keywords: Integral imaging, Point spread function, Simulation, Wave optics

DOIs:

10.1109/3DTV.2018.8478568

Bibliographical note

jufoid=50006

Source: Scopus

Source ID: 85056164335

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Understanding animals: A critical challenge in ACI

We present a qualitative content analysis of visual-verbal social media posts, where ordinary dog owners pretend to be their canine, to identify meaningful facets in their dogs' life-worlds, e.g. pleasures of human-dog relation, dog-dog relations, food etc. We use this knowledge to inform design of "quantified pets". The study targets a general problem in Animal-Computer Interaction (ACI), i.e. to understand animals when designing "for" them, although lacking a common language. Several approaches, e.g. ethnography and participatory design, have been appropriated from HCI without exhausting the issue. We argue for a methodological creativity and pluralism by suggesting an additional approach drawing on "kinesthetic empathy". It implies to understand animals by empathizing with their bodily movements over time and decoding the realities of their life-worlds. This, and other related approaches, has inspired animal researchers to conduct more or less radical participant observations during extensive duration to understand the perspective of the other. We suggest that dog owners whom share their lives with their dogs already possess a similar understanding as these experts, and thus uphold important experiences of canine life that could be used to understand individual dogs and inspire design.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Pervasive Computing, Research area: User experience, Stockholm University

Contributors: Aspling, F., Juhlin, O., Väättäjä, H.

Number of pages: 13

Pages: 148-160

Publication date: 29 Sep 2018

Host publication information

Title of host publication: NordiCHI 2018 : Revisiting the Life Cycle - Proceedings of the 10th Nordic Conference on Human-Computer Interaction

Publisher: ACM

ISBN (Electronic): 9781450364379

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Animal-Computer Interaction, Dog Blogs, Kinesthetic Empathy, Pet Dogs, Quantified Pets, Social Media

DOIs:

10.1145/3240167.3240226

Source: Scopus

Source ID: 85056568856

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Probabilistic saliency estimation

In this paper, we model the salient object detection problem under a probabilistic framework encoding the boundary connectivity saliency cue and smoothness constraints into an optimization problem. We show that this problem has a closed form global optimum solution, which estimates the salient object. We further show that along with the probabilistic framework, the proposed method also enjoys a wide range of interpretations, i.e. graph cut, diffusion maps and one-class classification. With an analysis according to these interpretations, we also find that our proposed method provides approximations to the global optimum to another criterion that integrates local/global contrast and large area saliency cues. The proposed unsupervised approach achieves mostly leading performance compared to the state-of-the-art unsupervised algorithms over a large set of salient object detection datasets including around 17k images for several evaluation metrics. Furthermore, the computational complexity of the proposed method is favorable/comparable to many state-of-the-art unsupervised techniques.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Signal Processing, Research group: Multimedia Research Group - MRG
Contributors: Aytekin, C., Iosifidis, A., Gabbouj, M.
Number of pages: 14
Pages: 359-372
Publication date: 2018
Peer-reviewed: Yes
Early online date: 20 Sep 2017

Publication information

Journal: Pattern Recognition
Volume: 74
ISSN (Print): 0031-3203
Ratings:
Scopus rating (2018): CiteScore 10.4 SJR 1.363 SNIP 3.211
Original language: English
ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Artificial Intelligence
Keywords: Diffusion maps, One-class classification, Probabilistic model, Saliency, Salient object detection, Spectral graph cut
DOIs:
10.1016/j.patcog.2017.09.023
Source: Scopus
Source ID: 85032271491
Research output: Contribution to journal › Article › Scientific › peer-review

Deep multiresolution color constancy

In this paper, a computational color constancy method is proposed via estimating the illuminant chromaticity in a scene by pooling from many local estimates. To this end, first, for each image in a dataset, we form an image pyramid consisting of several scales of the original image. Next, local patches of certain size are extracted from each scale in this image pyramid. Then, a convolutional neural network is trained to estimate the illuminant chromaticity per-patch. Finally, two more consecutive trainings are conducted, where the estimation is made per-image via taking the mean (1st training) and median (2nd training) of local estimates. The proposed method is shown to outperform the state-of-the-art in a widely used color constancy dataset.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Signal Processing, Intel Finland
Contributors: Aytekin, C., Nikkanen, J., Gabbouj, M.
Number of pages: 5
Pages: 3735-3739
Publication date: 20 Feb 2018

Host publication information

Title of host publication: 2017 IEEE International Conference on Image Processing, ICIP 2017 - Proceedings
Publisher: IEEE COMPUTER SOCIETY PRESS
ISBN (Electronic): 9781509021758
ASJC Scopus subject areas: Software, Computer Vision and Pattern Recognition, Signal Processing
Keywords: Color constancy, Deep learning, Illuminant chromaticity estimation, Local estimation, Multi-resolution
DOIs:
10.1109/ICIP.2017.8296980

Bibliographical note

jufoid=57423
Source: Scopus
Source ID: 85045299547
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Modeling and estimation of signal-dependent and correlated noise

The additive white Gaussian noise (AWGN) model is ubiquitous in signal processing. This model is often justified by central-limit theorem (CLT) arguments. However, whereas the CLT may support a Gaussian distribution for the random errors, it does not provide any justification for the assumed additivity and whiteness. As a matter of fact, data acquired in real applications can seldom be described with good approximation by the AWGN model, especially because errors are

typically correlated and not additive. Failure to model accurately the noise leads to inaccurate analysis, ineffective filtering, and distortion or even failure in the estimation. This chapter provides an introduction to both signal-dependent and correlated noise and to the relevant models and basic methods for the analysis and estimation of these types of noise. Generic one-parameter families of distributions are used as the essential mathematical setting for the observed signals. The distribution families covered as leading examples include Poisson, mixed Poisson–Gaussian, various forms of signal-dependent Gaussian noise (including multiplicative families and approximations of the Poisson family), as well as doubly censored heteroskedastic Gaussian distributions. We also consider various forms of noise correlation, encompassing pixel and readout cross-talk, fixed-pattern noise, column/row noise, etc., as well as related issues like photo-response and gain nonuniformity. The introduced models and methods are applicable to several important imaging scenarios and technologies, such as raw data from digital camera sensors, various types of radiation imaging relevant to security and to biomedical imaging.

General information

Publication status: Published

MoE publication type: A3 Part of a book or another research book

Organisations: Signal Processing, Research group: Signal and Image Restoration-RST, University of São Paulo

Contributors: Azzari, L., Borges, L. R., Foi, A.

Number of pages: 36

Pages: 1-36

Publication date: 2018

Host publication information

Title of host publication: Denoising of Photographic Images and Video : Fundamentals, Open Challenges and New Trends

Publisher: SPRINGER-VERLAG LONDON LTD

ISBN (Print): 978-3-319-96028-9

Publication series

Name: Advances in Computer Vision and Pattern Recognition

ISSN (Print): 2191-6586

ISSN (Electronic): 2191-6594

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Artificial Intelligence
Electronic versions:

Azzari2018-Modeling_and_estimation_of_signal_dependent_and_correlated_noise. Embargo ended: 11/09/19

DOIs:

10.1007/978-3-319-96029-6_1

URLs:

<http://urn.fi/URN:NBN:fi:tty-201907151959>. Embargo ended: 11/09/19

Source: Scopus

Source ID: 85053409603

Research output: Chapter in Book/Report/Conference proceeding › Chapter › Scientific › peer-review

Urban 3D segmentation and modelling from street view images and LiDAR point clouds

3D urban maps with semantic labels and metric information are not only essential for the next generation robots such as autonomous vehicles and city drones, but also help to visualize and augment local environment in mobile user applications. The machine vision challenge is to generate accurate urban maps from existing data with minimal manual annotation. In this work, we propose a novel methodology that takes GPS registered LiDAR (Light Detection And Ranging) point clouds and street view images as inputs and creates semantic labels for the 3D points clouds using a hybrid of rule-based parsing and learning-based labelling that combine point cloud and photometric features. The rule-based parsing boosts segmentation of simple and large structures such as street surfaces and building facades that span almost 75% of the point cloud data. For more complex structures, such as cars, trees and pedestrians, we adopt boosted decision trees that exploit both structure (LiDAR) and photometric (street view) features. We provide qualitative examples of our methodology in 3D visualization where we construct parametric graphical models from labelled data and in 2D image segmentation where 3D labels are back projected to the street view images. In quantitative evaluation we report classification accuracy and computing times and compare results to competing methods with three popular databases: NAVTEQ True, Paris-Rue-Madame and TLS (terrestrial laser scanned) Velodyne.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Signal Processing, Research group: Vision, Nokia

Contributors: Babahajani, P., Fan, L., Kämäräinen, J., Gabbouj, M.

Number of pages: 16

Pages: 679–694

Publication date: 2017

Peer-reviewed: Yes

Publication information

Journal: Machine Vision and Applications

Volume: 28

Issue number: 7

ISSN (Print): 0932-8092

Ratings:

Scopus rating (2017): CiteScore 5.3 SJR 0.485 SNIP 1.683

Original language: English

ASJC Scopus subject areas: Software, Hardware and Architecture, Computer Vision and Pattern Recognition, Computer Science Applications

Keywords: LiDAR, Point cloud, Robotics, Semantic segmentation, Street view, Urban 3D

Electronic versions:

Urban 3D segmentation and modelling from street view images and LiDAR point clouds

DOIs:

10.1007/s00138-017-0845-3

URLs:

<http://urn.fi/URN:NBN:fi:tty-201706121590>

Bibliographical note

EXT="Babahajiani, Pouria"

Source: Scopus

Source ID: 85019692066

Research output: Contribution to journal › Article › Scientific › peer-review

Effects of extrinsic noise are promoter kinetics dependent

Studies in *Escherichia coli* using in vivo single-RNA detection and time-lapse confocal microscopy showed that transcription is a multiple rate-limiting steps process, in agreement with previous in vitro measurements. Here, from simulations of a stochastic model of transcription validated empirically that accounts for cell-to-cell variability in RNA polymerase (RNAP) numbers, we investigate the hypothesis that the cell-to-cell variability in RNA numbers due to RNAP variability differs with the promoter rate-limiting steps dynamics. We find that increasing the cell-to-cell variability in RNAP numbers increases the cell-to-cell diversity in RNA numbers, but the degree with which it increases is promoter kinetics dependent. Namely, promoters whose open complex formation is relatively longer lasting dampen more efficiently this noise propagation phenomenon. We conclude that cell-to-cell variability in RNA numbers due to variability in RNAP numbers is promoter-sequence dependent and, thus, evolvable.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: BioMediTech, Faculty of Biomedical Sciences and Engineering, Research group: Laboratory of Biosystem Dynamics-LBD

Contributors: Bahrudeen, M. N., Startceva, S., Ribeiro, A. S.

Number of pages: 4

Pages: 44-47

Publication date: 14 May 2017

Host publication information

Title of host publication: Proceedings of the 2017 9th International Conference on Bioinformatics and Biomedical Technology, ICBBT 2017

Publisher: ACM

ISBN (Electronic): 9781450348799

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Extrinsic noise, Gene expression, Phenotypic diversity, Rate-limiting steps, Stochastic models, Transcription initiation

Electronic versions:

Effects_of_extrinsic_noise_are_promoter_kinetics_dependent

DOIs:

10.1145/3093293.3093295

URLs:

<http://urn.fi/URN:NBN:fi:tuni-201911186013>

Source: Scopus

Source ID: 85025117782

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Dynamics analysis of a redundant parallel manipulator driven by elastic cables

In this paper the dynamic analysis of a cable-driven parallel manipulator is studied in detail. The manipulator architecture is a simplified planar version adopted from the structure of Large Adaptive Reflector (LAR), the Canadian design of next generation giant radio telescopes. This structure consists of a parallel redundant manipulator actuated by long cables. The dynamic equations of this structure are nonlinear and implicit. Long cables, large amounts of impelling forces and high accelerations raise more concern about the elasticity of cables during dynamic analysis, which has been neglected in the preceding works. In this paper, the kinematic analysis of such manipulator is illustrated first. Then the nonlinear dynamic of such mechanism is derived using Newton-Euler formulation. Next a simple model for cable dynamics containing elastic and damping behavior is proposed. The proposed model neither ignores longitude elasticity properties of cable nor makes dynamic formulations heavily complicated like previous researches. Finally, manipulator dynamic with cable dynamic is derived, and the cable elasticity effects are compared in a simulation study. The results show significant role of elasticity in a cable-driven parallel manipulator such as the one used in LAR mechanism.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: K. N. Toosi University of Technology

Contributors: Bedoustani, Y. B., Taghirad, H. D., Aref, M. M.

Number of pages: 7

Pages: 536-542

Publication date: 2008

Host publication information

Title of host publication: 2008 10th International Conference on Control, Automation, Robotics and Vision, ICARCV 2008

Article number: 4795575

ISBN (Print): 9781424422876

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Control and Systems Engineering, Electrical and Electronic Engineering

Keywords: Cable-driven parallel manipulator, Dynamics, Elastic cable, Kinematics, Redundant manipulator

DOIs:

10.1109/ICARCV.2008.4795575

URLs:

<http://www.scopus.com/inward/record.url?scp=64549152409&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 64549152409

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Color game: A collaborative social robotic game for icebreaking; Towards the design of robotic ambiances as part of smart building services

Social robots are entering our workplaces, homes, medical and educational systems in assistive and collaborative roles. In our research, we have investigated the use of a social robot Pepper as an interactive icebreaker host to create a positive atmosphere at events. This paper presents two user studies (total n=43) in which we evaluated two interactive prototypes of playful applications on Pepper, with the overall aim of providing a personal and entertaining service for event attendees. Data about users' experiences and attitudes were collected with semi-structured interviews, surveys, and observations. The results of the studies suggest that the majority of the participants had pleasurable and positive experiences with the robot and its applications. Moreover, their positive encounters led them to accept social robots as icebreaker hosts to connect with strangers. Based on our findings, we present a list of design implications to help the future design of social robots used to facilitate social connectedness, and to aid in the development of social robots as intelligent agents performing tasks as integrated parts of smart spaces.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Civil Engineering, Research group: Digitalization in the real estate and construction sector

Contributors: Beheshtian, N., Kaipainen, K., Kähkönen, K., Ahtinen, A.

Number of pages: 10

Pages: 10-19

Publication date: 29 Jan 2020

Host publication information

Title of host publication: AcademicMindtrek 2020 - Proceedings of the 23rd International Academic Mindtrek Conference : January 2020, Tampere

Publisher: ACM

ISBN (Electronic): 9781450377744

ASJC Scopus subject areas: Software, Human-Computer Interaction, Computer Vision and Pattern Recognition, Computer Networks and Communications

Keywords: human robot interaction, ice breaking, smart building, social connectedness, social robots, user experience
DOIs:

10.1145/3377290.3377292

Source: Scopus

Source ID: 85080911326

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Compressive sensed video recovery via iterative thresholding with random transforms

The authors consider the problem of compressive sensed video recovery via iterative thresholding algorithm. Traditionally, it is assumed that some fixed sparsifying transform is applied at each iteration of the algorithm. In order to improve the recovery performance, at each iteration the thresholding could be applied for different transforms in order to obtain several estimates for each pixel. Then the resulting pixel value is computed based on obtained estimates using simple averaging. However, calculation of the estimates leads to significant increase in reconstruction complexity. Therefore, the authors propose a heuristic approach, where at each iteration only one transform is randomly selected from some set of transforms. First, they present simple examples, when block-based 2D discrete cosine transform is used as the sparsifying transform, and show that the random selection of the block size at each iteration significantly outperforms the case when fixed block size is used. Second, building on these simple examples, they apply the proposed approach when video block-matching and 3D filtering (VBM3D) is used for the thresholding and show that the random transform selection within VBM3D allows to improve the recovery performance as compared with the recovery based on VBM3D with fixed transform.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Computing Sciences, Research group: Computational Imaging-CI, ITMO University, Linköping University, University of Oulu

Contributors: Belyaev, E., Codreanu, M., Juntti, M., Egiazarian, K.

Number of pages: 14

Pages: 1187-1200

Publication date: 2020

Peer-reviewed: Yes

Publication information

Journal: IET Image Processing

Volume: 14

Issue number: 6

ISSN (Print): 1751-9659

Original language: English

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Electrical and Electronic Engineering

DOIs:

10.1049/iet-ipr.2019.0661

Bibliographical note

EXT="Belyaev, Evgeny"

Source: Scopus

Source ID: 85084175769

Research output: Contribution to journal › Article › Scientific › peer-review

An image guided treatment platform for prostate cancer photodynamic therapy

This study describes a multimodality images based platform to drive photodynamic therapies of prostate cancer using WST 11 TOOKAD Soluble drug. The platform integrates a pre-treatment planning tool based on magnetic resonance imaging and a per-treatment guidance tool based on transrectal ultrasound images. Evaluation of the platform on clinical data showed that prediction of the therapy outcome was possible with an accuracy of 90 %.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Frontier Photonics, Lille University Hospital - CHRU, Inserm

Contributors: Betrouni, N., Colin, P., Puech, P., Villers, A., Mordon, S.

Number of pages: 4
Pages: 370-373
Publication date: 2013

Host publication information

Title of host publication: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2013

Article number: 6609514

ISBN (Print): 9781457702167

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Signal Processing, Biomedical Engineering, Health Informatics

DOIs:

10.1109/EMBC.2013.6609514

URLs:

<http://www.scopus.com/inward/record.url?scp=84886469344&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84886469344

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Multi-factor authentication for wearables: Configuring system parameters with risk function

The users of today are already about to enter the era of highly integrated modern wearable devices—the time when smart accessories will, in turn, push aside regular Smartphones and Tablets bringing a variety of new security challenges. The number of simultaneously used bio-sensors, both integrated into smart wearables and connected over wireless interfaces, allows novel opportunities for Multi-factor Authentication (MFA) of the user. This manuscript proposes a solution for configuring the MFA based on the averaged direct and indirect losses risk analysis. The example application of Bayesian function for MFA presents the applicability of the proposed framework for the utilization with wearables.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Electronics and Communications Engineering, ITMO University, St. Petersburg State University of Aerospace Instrumentation

Contributors: Bezzateev, S., Afanasyeva, A., Voloshina, N., Ometov, A.

Publication date: 13 Nov 2017

Host publication information

Title of host publication: Proceedings of the 2nd International Conference on Advanced Wireless Information, Data, and Communication Technologies, AWICT 2017

Publisher: ACM

ISBN (Electronic): 9781450353106

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Information security, Multi-factor authentication, Risk function, Wearables

Electronic versions:

Multi-factor authentication for wearables 2017

DOIs:

10.1145/3231830.3231834

URLs:

<http://urn.fi/URN:NBN:fi:tuni-202002262382>

Source: Scopus

Source ID: 85045304145

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Data-driven stream mining systems for computer vision

In this chapter, we discuss the state of the art and future challenges in adaptive stream mining systems for computer vision. Adaptive stream mining in this context involves the extraction of knowledge from image and video streams in real-time, and from sources that are possibly distributed and heterogeneous. With advances in sensor and digital processing technologies, we are able to deploy networks involving large numbers of cameras that acquire increasing volumes of image data for diverse applications in monitoring and surveillance. However, to exploit the potential of such extensive networks for image acquisition, important challenges must be addressed in efficient communication and analysis of such data under constraints on power consumption, communication bandwidth, and end-to-end latency. We discuss these challenges in this chapter, and we also discuss important directions for research in addressing such challenges using dynamic, data-driven methodologies.

General information

Publication status: Published
MoE publication type: A3 Part of a book or another research book
Organisations: Signal Processing Research Community (SPRC), University of Maryland, Electrical Engineering Department, University of California, Los Angeles (UCLA)
Contributors: Bhattacharyya, S. S., Van Der Schaar, M., Atan, O., Tekin, C., Sudusinghe, K.
Number of pages: 16
Pages: 249-264
Publication date: 2014

Host publication information

Title of host publication: Advances in Computer Vision and Pattern Recognition
Volume: 68
Publisher: SPRINGER-VERLAG LONDON LTD

Publication series

Name: Advances in Computer Vision and Pattern Recognition
Volume: 68
ISSN (Print): 21916586
ISSN (Electronic): 21916594
ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Artificial Intelligence
DOIs:
10.1007/978-3-319-09387-1_12
URLs:
<http://www.scopus.com/inward/record.url?scp=84984919867&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84984919867
Research output: Chapter in Book/Report/Conference proceeding > Chapter > Scientific > peer-review

Application-specific instruction processor for extracting local binary patterns

Local Binary Pattern (LBP) is texture operator used in preprocessing for object detection, tracking, face recognition and fingerprint matching. Many of these applications are performed on embedded devices, which poses limitations on the implementation complexity and power consumption. As LBP features are computed pixelwise, high performance is required for real time extraction of LBP features from high resolution video. This paper presents an application-specific instruction processor for LBP extraction. The compact, yet powerful processor is capable of extracting LBP features from 1280 × 720p (30 fps) video with a reasonable 304 MHz clock rate. With a low power consumption and an area of less than 16k gates the processor is suitable for embedded devices. Experiments present resource and power consumption measured on an FPGA board, along with processor synthesis results. In terms of latency, our processor requires 17.5 × less clock cycles per LBP feature than a workstation implementation and only 2.0 × more than a hardwired ASIC.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Signal Processing Research Community (SPRC), Dept. of Computer Science and Engineering, Univ of Oulu
Contributors: Boutellier, J., Lundbom, I., Janhunen, J., Ylimainen, J., Hannuksela, J.
Number of pages: 8
Pages: 82-89
Publication date: 2012

Host publication information

Title of host publication: DASIP 2012 - Proceedings of the 2012 Conference on Design and Architectures for Signal and Image Processing
Article number: 6385363
ISBN (Print): 9782953998726
ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Hardware and Architecture, Signal Processing, Electrical and Electronic Engineering
Keywords: Digital signal processors, Feature extraction, Image texture analysis, Video signal processing
URLs:
<http://www.scopus.com/inward/record.url?scp=84872397244&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84872397244
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Improved weighted prediction based color gamut scalability in SHVC

One use case that the scalable extension (SHVC) of the state-of-the-art High Efficiency Video Coding (HEVC) standard aims for is to support Ultra High Definition (UHD) TV broadcast in a backwards compatible way with the existing High Definition (HD) TV broadcast. However, since UHD content typically has higher bit-depth and wider color gamut in addition to increased spatial resolution, the compression efficiency is highly affected by the inter-layer processing applied on the base layer picture. This paper proposes an improvement for the weighted prediction based color gamut scalability to have a better mapping between the color gamuts of the base and enhancement layers. The proposed method aims at capturing the nonlinear characteristics of the color gamut mapping using a piecewise linear model, whose parameters are signaled through weighted prediction mechanism and multiple inter-layer reference pictures. Compared to other existing methods for color gamut mapping in SHVC, such as the 3D Look Up Table (LUT) method, the proposed weighted prediction based approach is less complex, as it does not require any changes to the decoder. The simulation results show up to 3.8% Bjontegaard delta bitrate gain in luma for all intra and 3.0% for random access configurations compared to the existing weighted prediction based scalability method in SHVC.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: Video, Research Community on Data-to-Decision (D2D), Åbo Akademi, Department of Information Technologies, Nokia, AVCR Information Technologies

Contributors: Bugdayci Sansli, D., Aminlou, A., Ugur, K., Hannuksela, M. M., Gabbouj, M.

Number of pages: 4

Pages: 201-204

Publication date: 27 Feb 2015

Host publication information

Title of host publication: 2014 IEEE Visual Communications and Image Processing Conference, VCIP 2014

Publisher: The Institute of Electrical and Electronics Engineers, Inc.

Article number: 7051539

ISBN (Print): 9781479961399

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition

Keywords: BT.2020, BT.709, HEVC, scalable video coding, SHVC, weighted prediction, wide color gamut

DOIs:

10.1109/VCIP.2014.7051539

Bibliographical note

EXT="Ugur, Kemal"

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Satisfaction and willingness to consume immersive journalism: Experiment of differences between VR, 360 video, and article

Immersive journalism has been touted to revolutionize journalism due to its ability to afford a multi-modal engrossing experience. However, hardly any experiments have been conducted whether consumers' satisfaction and consequent intentions to use immersive journalistic media may differ from traditional forms of journalistic content. Therefore, in this study, we investigate the differences in satisfaction and continued use intentions between article, 360 video and VR-based interaction with content. The data was collected via a randomized controlled laboratory experiment with between-subjects design (N = 87). Participants were randomly assigned to reading a written article based on the video (article) and watching the video on a computer screen (2D 360) or in mobile VR (VR 360). The collected data consisted of demographics (age and gender) and reported satisfaction and intention to continue use. Results suggest that those who were assigned to VR 360 had higher intentions to continue use, but not greater satisfaction than those in the other two conditions. However, the intention was predicted to an extent by satisfaction as suggested by previous literature. Finally, age and gender did not predict continued use. These findings imply that users prefer the new media technology for consuming journalism content and support previous findings of the relationship between satisfaction and intention to continue use. Finally, avenues for further research are presented.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Research group: TUT Game Lab

Contributors: Bujic, M., Hamari, J.

Number of pages: 6

Pages: 120-125

Publication date: 29 Jan 2020

Host publication information

Title of host publication: AcademicMindtrek 2020 - Proceedings of the 23rd International Academic Mindtrek Conference : January 2020, Tampere

Publisher: ACM

ISBN (Electronic): 9781450377744

ASJC Scopus subject areas: Software, Human-Computer Interaction, Computer Vision and Pattern Recognition, Computer Networks and Communications

Keywords: 360-degree video, age, gender, immersive journalism, intention to continue use, satisfaction, virtual reality
DOIs:

10.1145/3377290.3377310

Source: Scopus

Source ID: 85080895604

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

The Finnish you – An interactive storytelling application for an airport environment

Traveling should be full of excitement and new experiences. However, a chaotic airport environment and constant waiting often halt these pleasurable feelings. Although passengers can spend their time shopping, they are unlikely to connect personally to the products. Furthermore, airport services seldom highlight the local culture that passengers miss by being confined to the airport. To address these shortcomings, we present a mobile web-application, called “The Finnish You”. Utilizing the elements of interactive storytelling and gamification, the application guides users through shops and brands in the airport while teaching about the local culture in a personalized way. The application was tested in a user study with nine participants in a controlled office environment and was seen as a satisfactory way to spend time waiting in the airport. Our findings show how a personalized storytelling approach may convert ordinary shopping activity into a culture-learning adventure. We further suggest implications for the design of storytelling applications regarding the airport context of use.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Faculty of Computing and Electrical Engineering, Pervasive Computing, Research area: User experience, Human-Centered Technology (IHTE)

Contributors: Burova, A., Kelling, C., Hakulinen, J., Kallioniemi, P., Keskinen, T., Turunen, M., Väättäjä, H.

Number of pages: 10

Pages: 182-191

Publication date: 10 Oct 2018

Host publication information

Title of host publication: Mindtrek 2018 - Proceedings of the 22nd International Academic Mindtrek Conference

Publisher: ACM

ISBN (Electronic): 9781450365895

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Airport Environment, Digital Storytelling, Gamification, Mobile Web-Application, User Experience

DOIs:

10.1145/3275116.3275142

Source: Scopus

Source ID: 85056694022

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Promoting local culture and enriching airport experiences through interactive storytelling

Experiences in airports may shape future travel plans and contribute to tourism destination development. However, a chaotic environment and time-consuming procedural routines in airports may result in negative associations towards the host country and its culture. Despite the existence of assistive airport applications, little attention is given to facilitating travelers' engagement with cultural exploration. This paper introduces a concept of interactive personalized storytelling that provides both a cultural learning adventure and connection to local retailing. Our application generates an imaginative Finnish storyline unique to every user to guide them through local shops in the airport. A field evaluation was conducted with 15 travelers of different nationalities. Travelers perceived the interactive storytelling experience as an interesting and unique way to spend waiting time at the airport while increasing cultural exposure. Moreover, we found this method to be effective in persuading travelers to explore local products at the airport. Further, our results give insight to designing storytelling applications for large public places.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Lapland University of Applied Sciences, Tampere University, Tampere University, Tampere University

Contributors: Burova, A., Kelling, C., Keskinen, T., Hakulinen, J., Kallioniemi, P., Väättäjä, H., Turunen, M.

Number of pages: 7

Publication date: 26 Nov 2019

Host publication information

Title of host publication: MUM 2019 - 18th International Conference on Mobile and Ubiquitous Multimedia, Proceedings
Publisher: Association for Computing Machinery
Editors: Jacucci, G., Paterno, F., Rohs, M., Santoro, C.
Article number: 3365640
ISBN (Electronic): 9781450376242

Publication series

Name: ACM International Conference Proceeding Series
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Airport experience, Digital storytelling, Field study, Mobile application
DOIs:
10.1145/3365610.3365640
URLs:
<http://urn.fi/URN:NBN:fi:tuni-202001101166>

Bibliographical note

EXT="Vääätäjä, Heli"
Source: Scopus
Source ID: 85076809996
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Convolutional low-resolution fine-grained classification

Successful fine-grained image classification methods learn subtle details between visually similar (sub-)classes, but the problem becomes significantly more challenging if the details are missing due to low resolution. Encouraged by the recent success of Convolutional Neural Network (CNN) architectures in image classification, we propose a novel resolution-aware deep model which combines convolutional image super-resolution and convolutional fine-grained classification into a single model in an end-to-end manner. Extensive experiments on multiple benchmarks demonstrate that the proposed model consistently performs better than conventional convolutional networks on classifying fine-grained object classes in low-resolution images.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Signal Processing, Research group: Vision
Contributors: Cai, D., Chen, K., Qian, Y., Kämäräinen, J.
Pages: 166-171
Publication date: Mar 2019
Peer-reviewed: Yes
Early online date: 2017

Publication information

Journal: Pattern Recognition Letters
Volume: 119
ISSN (Print): 0167-8655
Ratings:
Scopus rating (2019): CiteScore 6.3 SJR 0.848 SNIP 2.021
Original language: English
ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Artificial Intelligence
Keywords: Deep learning, Fine-grained image classification, Super resolution convolutional neural networks
DOIs:
10.1016/j.patrec.2017.10.020
Source: Scopus
Source ID: 85032974725
Research output: Contribution to journal > Article > Scientific > peer-review

Design space exploration and implementation of RVC-CAL applications using the TURNUS framework

While research on the design of heterogeneous concurrent systems has a long and rich history, a unified design methodology and tool support has not emerged so far, and thus the creation of such systems remains a difficult, time-consuming and error-prone process. The absence of principled support for system evaluation and optimization at high abstraction levels makes the quality of the resulting implementation highly dependent on the experience or prejudices of the designer. In this work we present TURNUS, a unified dataflow design space exploration framework for heterogeneous

parallel systems. It provides high-level modelling and simulation methods and tools for system level performances estimation and optimization. TURNUS represents the outcome of several years of research in the area of co-design exploration for multimedia stream applications. During the presentation, it will be demonstrated how the initial high-level abstraction of the design facilitates the use of different analysis and optimization heuristics. These guide the designer during validation and optimization stages without requiring low-level implementations of parts of the application. Our framework currently yields exploration and optimization results in terms of algorithmic optimization, rapid performance estimation, application throughput, buffer size dimensioning, and power optimization.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing Research Community (SPRC), CRPP, Lund University, Dept. of Computer Science and Engineering, Univ of Oulu

Contributors: Casale-Brunet, S., Bezati, E., Alberti, C., Roquier, G., Mattavelli, M., Janneck, J. W., Boutellier, J.

Number of pages: 2

Pages: 341-342

Publication date: 2013

Host publication information

Title of host publication: DASIP 2013 - Proceedings of the 2013 Conference on Design and Architectures for Signal and Image Processing

Article number: 6661566

ISBN (Print): 9791092279016

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Hardware and Architecture, Signal Processing, Electrical and Electronic Engineering

Keywords: Co-exploration, Dataflow, Design space

URLs:

<http://www.scopus.com/inward/record.url?scp=84892650917&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84892650917

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

A Primal Neural Network for Online Equality-Constrained Quadratic Programming

This paper aims at solving online equality-constrained quadratic programming problem, which is widely encountered in science and engineering, e.g., computer vision and pattern recognition, digital signal processing, and robotics. Recurrent neural networks such as conventional GradientNet and ZhangNet are considered as powerful solvers for such a problem in light of its high computational efficiency and capability of circuit realisation. In this paper, an improved primal recurrent neural network and its electronic implementation are proposed and analysed. Compared to the existing recurrent networks, i.e. GradientNet and ZhangNet, our network can theoretically guarantee superior global exponential convergence. Robustness performance of our such neural model is also analysed under a large model implementation error, with the upper bound of steady-state solution error estimated. Simulation results demonstrate theoretical analysis on the proposed model, which also verify the effectiveness of the proposed model for online equality-constrained quadratic programming.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Signal Processing, Research group: Vision, Shanghai Institute of Ceramics Chinese Academy of Sciences, Institute of Automation Chinese Academy of Sciences

Contributors: Chen, K., Zhang, Z.

Number of pages: 8

Pages: 381-388

Publication date: 2018

Peer-reviewed: Yes

Publication information

Journal: Cognitive Computation

Volume: 10

Issue number: 2

ISSN (Print): 1866-9956

Ratings:

Scopus rating (2018): CiteScore 7.1 SJR 1.06 SNIP 1.965

Original language: English

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Science Applications, Cognitive Neuroscience

Keywords: Global exponential convergence, Online equality-constrained quadratic programming, Recurrent neural networks, Robustness analysis

DOIs:

10.1007/s12559-017-9510-4

Source: Scopus

Source ID: 85030320446

Research output: Contribution to journal › Article › Scientific › peer-review

"The superhero of the university": Experience-driven design and field study of the university guidance robot

Robots have recently gained popularity in customer service. Especially social robots are nowadays utilized in healthcare, elderly homes and schools. Although it is crucial to design social robots according to well-defined user experience goals, research related to experience-driven design of social robots is still scarce. Experience-Driven Design (EDD) is a framework to design interaction for technology based on certain goals, known as experience goals. In this paper, we present the design and evaluation of the university guidance robot based on the user experience goals defined in previous research. The experience goals are nurture, fellowship and recreation. We designed applications, interaction, and robot's behavior to support the fulfillment of the experience goals. The social robot Pepper served as a platform for the university guidance robot. The evaluation was conducted as a field study in a university campus with 32 university students during the orientation week. According to our findings, the university guide robot successfully evoked nurture, fellowship and recreation among participants.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences

Contributors: Chowdhury, A., Ahtinen, A., Kaipainen, K.

Number of pages: 9

Pages: 1-9

Publication date: 29 Jan 2020

Host publication information

Title of host publication: AcademicMindtrek 2020 - Proceedings of the 23rd International Academic Mindtrek Conference : January 2020, Tampere

Publisher: ACM

ISBN (Electronic): 9781450377744

ASJC Scopus subject areas: Software, Human-Computer Interaction, Computer Vision and Pattern Recognition, Computer Networks and Communications

Keywords: experience-driven design, social robots, user experience

DOIs:

10.1145/3377290.3377304

URLs:

<http://www.scopus.com/inward/record.url?scp=85080943314&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 85080943314

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Salient event detection in basketball mobile videos

Modern smartphones have become the most popular means for recording videos. In fact, thanks to their portability, smartphones allow for recording anything and at any moment of our everyday life. One common occasion is represented by sport happenings, where people often record their favourite team or players. Automatic analysis of such videos is important for enabling applications such as automatic organization, browsing and summarization of the content. This paper proposes novel algorithms for the detection of salient events in videos recorded at basketball games. The novel approach consists of jointly analyzing visual data and magnetometer data. The magnetometer data provides information about the horizontal orientation of the camera. The proposed joint analysis allows for a reduced number of false positives and for a reduced computational complexity. The algorithms are tested on data captured during real basketball games. The experimental results clearly show the advantages of the proposed approach.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: Video, Research Community on Data-to-Decision (D2D), Nokia Technologies

Contributors: Cricri, F., Mate, S., Curcio, I. D. D., Gabbouj, M.

Number of pages: 8

Pages: 63-70

Publication date: 5 Feb 2015

Host publication information

Title of host publication: Proceedings - 2014 IEEE International Symposium on Multimedia, ISM 2014

Publisher: The Institute of Electrical and Electronics Engineers, Inc.

Article number: 7032995

ISBN (Print): 978-1-4799-4312-8

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Human-Computer Interaction, Software, Media Technology

Keywords: Basketball, detection, event, mobile, video

DOIs:

10.1109/ISM.2014.67

Bibliographical note

EXT="Curcio, Igor D D"

EXT="Mate, Sujeet"

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Bandwidth reduction of omnidirectional viewport-dependent video streaming via subjective quality assessment

Omnidirectional video is more and more widespread in consumer electronics and professional capture devices, as well over Internet via novel streaming services. Omnidirectional video requires a large streaming bandwidth. To date there is little knowledge about the subjective experience of streaming services of omnidirectional video. The aim of this paper is to present subjective assessment results of experiments using a tilebased streaming system for omnidirectional video with the goal of reducing the streaming bandwidth. The results we present show that it is possible to reduce streaming bit rates by an average of 44% for a subjective DMOS value of 4.5 for different content genres.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Digital Media Laboratory, Nokia

Contributors: Curcio, I. D., Toukoma, H., Naik, D.

Number of pages: 6

Pages: 9-14

Publication date: 27 Oct 2017

Host publication information

Title of host publication: AltMM 2017 - Proceedings of the 2nd International Workshop on Multimedia Alternate Realities, co-located with MM 2017

Publisher: ACM

ISBN (Electronic): 9781450355070

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Human-Computer Interaction

Keywords: 360 degrees video, Omnidirectional video, Streaming bandwidth optimization, Subjective assessment, Subjective quality evaluation, Virtual reality streaming

DOIs:

10.1145/3132361.3132364

Bibliographical note

EXT="Curcio, Igor D.D."

INT="sgn,"Naik, Deepa"

Source: Scopus

Source ID: 85036610778

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Subjective evaluation of Super Multi-View compressed contents on high-end light-field 3D displays

Super Multi-View (SMV) video content is composed of tens or hundreds of views that provide a light-field representation of a scene. This representation allows a glass-free visualization and eliminates many causes of discomfort existing in current available 3D video technologies. Efficient video compression of SMV content is a key factor for enabling future 3D video services. This paper first compares several coding configurations for SMV content and several inter-view prediction structures are also tested and compared. The experiments mainly suggest that large differences in coding efficiency can be observed from one configuration to another. Several ratios for the number of coded and synthesized views are compared, both objectively and subjectively. It is reported that view synthesis significantly affects the coding scheme. The amount of views to skip highly depends on the sequence and on the quality of the associated depth maps. Reported ranges of bitrates required to obtain a good quality for the tested SMV content are realistic and coherent with future 4. K/8. K needs. The reliability of the PSNR metric for SMV content is also studied. Objective and subjective results show that PSNR is able to reflect increase or decrease in subjective quality even in the presence of synthesized views. However,

depending on the ratio of coded and synthesized views, the order of magnitude of the effective quality variation is biased by PSNR. Results indicate that PSNR is less tolerant to view synthesis artifacts than human viewers. Finally, preliminary observations are initiated. First, the light-field conversion step does not seem to alter the objective results for compression. Secondly, the motion parallax does not seem to be impacted by specific compression artifacts. The perception of the motion parallax is only altered by variations of the typical compression artifacts along the viewing angle, in cases where the subjective image quality is already low. To the best of our knowledge, this paper is the first to carry out subjective experiments and to report results of SMV compression for light-field 3D displays. It provides first results showing that improvement of compression efficiency is required, as well as depth estimation and view synthesis algorithms improvement, but that the use of SMV appears realistic according to next generation compression technology requirements.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Orange Labs, Holografika Kft., Pámany Peter Katolikus Egyetem

Contributors: Dricot, A., Jung, J., Cagnazzo, M., Pesquet, B., Dufaux, F., Kovács, P., Adhikarla, V. K.

Pages: 369–385

Publication date: 2015

Peer-reviewed: Yes

Publication information

Journal: Signal Processing: Image Communication

Volume: 39

Issue number: Part B

ISSN (Print): 0923-5965

Ratings:

Scopus rating (2015): CiteScore 4 SJR 0.532 SNIP 1.413

Original language: English

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Signal Processing, Software, Electrical and Electronic Engineering

Keywords: 3D, Light-field, Subjective evaluation, Super Multi-View, Video coding, Video compression

DOIs:

10.1016/j.image.2015.04.012

Source: Scopus

Source ID: 84947865722

Research output: [Contribution to journal](#) › [Article](#) › [Scientific](#) › [peer-review](#)

Two-time coherence of pulse trains and the integrated degree of temporal coherence

We examine the temporal coherence properties of trains of nonidentical short optical pulses in the framework of the second-order coherence theory of nonstationary light. Considering Michelson's interferometric measurement of temporal coherence, we demonstrate that time-resolved interferograms reveal the full two-time temporal coherence function of the partially coherent pulse train. We also show that the result given by the time-integrated Michelson interferogram equals the true degree of temporal coherence only when the pulse train is quasistationary, i.e., the coherence time is a small fraction of the pulse duration. True two-time and integrated coherence functions produced by specific models representing perturbed trains of mode-locked pulses and supercontinuum pulse trains produced in nonlinear fibers are illustrated.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Physics, Research group: Nonlinear Fiber Optics, Research area: Optics, Frontier Photonics, Itä-Suomen yliopisto, Institute of Photonics

Contributors: Dutta, R., Friberg, A. T., Genty, G., Turunen, J.

Number of pages: 7

Pages: 1631-1637

Publication date: 2015

Peer-reviewed: Yes

Publication information

Journal: Journal of the Optical Society of America A: Optics Image Science and Vision

Volume: 32

Issue number: 9

ISSN (Print): 1084-7529

Ratings:

Scopus rating (2015): CiteScore 3.4 SJR 0.918 SNIP 1.103

Original language: English

ASJC Scopus subject areas: Atomic and Molecular Physics, and Optics, Electronic, Optical and Magnetic Materials, Computer Vision and Pattern Recognition

DOIs:

10.1364/JOSAA.32.001631

URLs:

<http://www.scopus.com/inward/record.url?scp=84943414423&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84943414423

Research output: Contribution to journal › Article › Scientific › peer-review

Rate-distortion based reversible watermarking for JPEG images with quality factors selection

The improved reversible data hiding scheme which is a part of JPEG coding process is introduced. Generally, one of the common constraints imposed on digital watermarking in frequency domain is a small payload that can be embedded without causing high degradation of a JPEG stego image. Moreover, even at small hidden payload the stego image file size will increase to some extent. In no existing data hiding technique compliant with JPEG there is a possibility to define in advance the file size of the watermarked image. Therefore, in this paper we propose to use a rate-distortion theory that minimizes coding distortion subject to a coding rate constraint. An iterative algorithm based on a Lagrangian formulation is applied to obtain a vector of quality factors for each of the 8×8 blocks that scale the JPEG standard quantization table. The experimental results show the advantage of the proposed improved watermarking scheme in terms of data payload versus quality and file size compared with the state-of-the-art data hiding schemes, and, furthermore, clarify the improvements of its optimized counterpart. © 2013 University Paris 13.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Tampere University of Technology, Department of Signal Processing, Research Community on Data-to-Decision (D2D), Signal Processing Research Community (SPRC)

Contributors: Efimushkina, T., Egiazarian, K., Gabbouj, M.

Number of pages: 6

Pages: 94-99

Publication date: 2013

Host publication information

Title of host publication: 2013 4th European Workshop on Visual Information Processing, EUVIP 2013, Paris, France, 10.-12.2013

Publisher: University of Paris 13

Article number: 6623958

ISBN (Print): 978-82-93269-13-7

Publication series

Name: European Workshop on Visual Information Processing

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Signal Processing

Keywords: JPEG, Lagrangian relaxation, RateDistortion Optimization, reversible, watermarking

URLs:

<http://www.scopus.com/inward/record.url?scp=84890525586&partnerID=8YFLogxK> (Link to publication in Scopus)

Bibliographical note

Contribution: organisation=sgn,FACT1=1
Portfolio EDEND: 2013-12-29
Publisher name: University of Paris 13

Source: researchoutputwizard

Source ID: 2099

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Patterns for controlling chaos in a startup

A growing trend in industrial software engineering is that new software products and information services are developed under conditions of notable uncertainty. This is especially visible in startup enterprises which aim at new kinds of products and services in rapidly changing social web, where potential customers can quickly adopt new behavior. Special characteristics of the startups are lack of resources and funds, and startups may need to change direction fast. All these affect the software engineering practices used in the startups. Unfortunately almost 90 percent of all startups fail and goes bankrupt. There are probably indefinite numbers of reasons why startups fail. Failure might be caused by wrongly chosen software engineering practices or inconsiderate decision making. While there is no recipe for success, we argue that good practices that can help on the way to success can be identified from successful startups. In this paper, we present two patterns that startups can consider when entering the growth phase of the lifecycle.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Department of Pervasive Computing
Contributors: Eloranta, V.
Number of pages: 8
Pages: 1-8
Publication date: 10 Apr 2014

Host publication information

Title of host publication: VikingPLoP 2014 Proceedings of the 8th Nordic Conference on Pattern Languages of Programs
Volume: 2014-April
Publisher: Association for Computing Machinery
ISBN (Print): 9781450326605

Publication series

Name: ACM International Conference Proceeding Series
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Lean start-up, Organizational patterns, Patterns, Software engineering, Start-up
DOIs:
10.1145/2676680.2676682
Source: Scopus
Source ID: 84940028558
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

360 panorama super-resolution using deep convolutional networks

We propose deep convolutional neural network (CNN) based super-resolution for 360 (equirectangular) panorama images used by virtual reality (VR) display devices (e.g. VR glasses). Proposed super-resolution adopts the recent CNN architecture proposed in (Dong et al., 2016) and adapts it for equirectangular panorama images which have specific characteristics as compared to standard cameras (e.g. projection distortions). We demonstrate how adaptation can be performed by optimizing the trained network input size and fine-tuning the network parameters. In our experiments with 360 panorama images of rich natural content CNN based super-resolution achieves average PSNR improvement of 1.36 dB over the baseline (bicubic interpolation) and 1.56 dB by our equirectangular specific adaptation.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Signal Processing, Nokia Technologies
Contributors: Fakour-Sevom, V., Guldogan, E., Kämäräinen, J.
Number of pages: 7
Pages: 159-165
Publication date: 2018

Host publication information

Title of host publication: VISIGRAPP 2018 - Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications
Volume: 4
Publisher: SCITEPRESS
ISBN (Electronic): 9789897582905
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Graphics and Computer-Aided Design, Artificial Intelligence
Keywords: Deep convolutional neural network, Equirectangular panorama, Super-resolution, Virtual reality
DOIs:
10.5220/0006618901590165

Bibliographical note

EXT="Guldogan, Esin"
Source: Scopus
Source ID: 85047846712
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Foveated Nonlocal Self-Similarity

When we gaze a scene, our visual acuity is maximal at the fixation point (imaged by the fovea, the central part of the retina) and decreases rapidly towards the periphery of the visual field. This phenomenon is known as foveation. We investigate the role of foveation in nonlocal image filtering, installing a different form of self-similarity: the foveated self-similarity. We consider the image denoising problem as a simple means of assessing the effectiveness of descriptive

models for natural images and we show that, in nonlocal image filtering, the foveated self-similarity is far more effective than the conventional windowed self-similarity. To facilitate the use of foveation in nonlocal imaging algorithms, we develop a general framework for designing foveation operators for patches by means of spatially variant blur. Within this framework, we construct several parametrized families of operators, including anisotropic ones. Strikingly, the foveation operators enabling the best denoising performance are the radial ones, in complete agreement with the orientation preference of the human visual system.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research area: Signal and Information Processing, Research group: Signal and Image Restoration-RST, Dipartimento di Elettronica, Politecnico di Milano

Contributors: Foi, A., Boracchi, G.

Number of pages: 33

Pages: 78–110

Publication date: 2016

Peer-reviewed: Yes

Early online date: 9 Mar 2016

Publication information

Journal: International Journal of Computer Vision

Volume: 120

Issue number: 1

ISSN (Print): 0920-5691

Ratings:

Scopus rating (2016): CiteScore 17.4 SJR 6.779 SNIP 5.171

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition

Electronic versions:

Foveated Nonlocal Self-Similarity. Embargo ended: 9/03/17

DOIs:

10.1007/s11263-016-0898-1

URLs:

<http://urn.fi/URN:NBN:fi:tty-201605274192>. Embargo ended: 9/03/17

Bibliographical note

EXT="Boracchi, Giacomo"

Source: Scopus

Source ID: 84960153979

Research output: Contribution to journal › Article › Scientific › peer-review

Performance analysis of single-query 6-DoF camera pose estimation in self-driving setups

In this work, we consider the problem of single-query 6-DoF camera pose estimation, i.e. estimating the position and orientation of a camera by using reference images and a point cloud. We perform a systematic comparison of three state-of-the-art strategies for 6-DoF camera pose estimation: feature-based, photometric-based and mutual-information-based approaches. Two standard datasets with self-driving setups are used for experiments, and the performance of the studied methods is evaluated in terms of success rate, translation error and maximum orientation error. Building on the analysis of the results, we evaluate a hybrid approach that combines feature-based and mutual-information-based pose estimation methods to benefit from their complementary properties for pose estimation. Experiments show that (1) in cases with large appearance change between query and reference, the hybrid approach outperforms feature-based and mutual-information-based approaches by an average increment of 9.4% and 8.7% in the success rate, respectively; (2) in cases where query and reference images are captured at similar imaging conditions, the hybrid approach performs similarly as the feature-based approach, but outperforms both photometric-based and mutual-information-based approaches with a clear margin; (3) the feature-based approach is consistently more accurate than mutual-information-based and photometric-based approaches when at least 4 consistent matching points are found between the query and reference images.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Computing Sciences, Universidad Industrial de Santander, Czech Technical University in Prague

Contributors: Fu, J., Pertuz, S., Matas, J., Kämäräinen, J.

Pages: 58-73

Publication date: Sep 2019

Peer-reviewed: Yes

Publication information

Journal: Computer Vision and Image Understanding

Volume: 186

ISSN (Print): 1077-3142

Ratings:

Scopus rating (2019): CiteScore 8.7 SJR 1.453 SNIP 2.255

Original language: English

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition

Keywords: 3D point cloud, Camera pose estimation, Hybrid method, Mutual information, Photometric matching, Self driving car

DOIs:

10.1016/j.cviu.2019.04.009

Bibliographical note

EXT="Matas, Jiri"

Source: Scopus

Source ID: 85067195521

Research output: Contribution to journal > Article > Scientific > peer-review

Light field reconstruction using shearlet transform in tensorflow

Shearlet Transform (ST) is one of the most effective approaches for light field reconstruction from Sparsely-Sampled Light Fields (SSLFs). This demo paper presents a comprehensive implementation of ST for light field reconstruction using one of the most popular machine learning libraries, i.e. Tensor Flow. The flexible architecture of TensorFlow allows for the easy deployment of ST across different platforms (CPUs, GPUs, TPUs) running varying operating systems with high efficiency and accuracy.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Research group: 3D MEDIA, Computer Science Institute

Contributors: Gao, Y., Koch, R., Bregovic, R., Gotchev, A.

Publication date: 1 Jul 2019

Host publication information

Title of host publication: 2019 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2019

Publisher: IEEE

ISBN (Electronic): 9781538692141

ASJC Scopus subject areas: Media Technology, Computer Vision and Pattern Recognition

Keywords: Epipolar-Plane Image, Light Field Reconstruction, Light Field Sparsification, Shearlet Transform, TensorFlow

Electronic versions:

icme2019w_253-submitted

DOIs:

10.1109/ICMEW.2019.00116

URLs:

<http://urn.fi/URN:NBN:fi:tuni-202001211439>

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Parametric exploration of cellular swelling in a computational model of cortical spreading depression

Cortical spreading depression (CSD) is a slowly propagating wave of depolarization of brain cells, followed by temporary silenced electrical brain activity. Major structural changes during CSD are linked to neuronal and possibly glial swelling. However, basic questions still remain unanswered. In particular, there are open questions regarding whether neurons or glial cells swell more, and how the cellular swelling affects the CSD wave propagation. In this study, we computationally explore how different parameters affect the swelling of neurons and astrocytes (starshaped glial cells) during CSD and how the cell swelling alters the CSD wave spatial distribution. We apply a homogenized mathematical model that describes electrodiffusion in the intraand extracellular space, and discretize the equations using a finite element method. The simulations are run with a twocompartment (extracellular space and neurons) and a threecompartment version of the model with astrocytes added. We consider cell swelling during CSD in four scenarios: (A) incorporating aquaporin-4 channels in the astrocytic membrane, (B) increasing the neuron/astrocyte ratio to 2:1, (C) blocking and increasing the Na^+/K^+ -ATPase rate in the astrocytic compartment, and (D) blocking the Cl^- channels in astrocytes. Our results show that increasing the water permeability in the astrocytes results in a higher astrocytic swelling and a lower neuronal swelling than in the default case. Further, elevated neuronal density increases the swelling in both neurons and astrocytes. Blocking the Na^+/K^+ -ATPase in the astrocytes leads to an increased wave width and swelling in both compartments, which instead decreases when the pump rate is raised. Blocking the Cl^- channels in the astrocytes results in neuronal swelling, and a shrinkage in the astrocytes. Our results suggest a supporting role of astrocytes in preventing cellular swelling and CSD, as well as highlighting how dysfunctions in astrocytes might elicit CSD.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: BioMediTech, Research group: Computational Biophysics and Imaging Group, Centre for Molecular Medicine Norway, Nordic European Molecular Biology Laboratory Partnership, University of Oslo, University of California San Diego, Simula Research Laboratory

Contributors: Genocchi, B., Cunha, A., Jain, S., Hyttinen, J., Lenk, K., Ellingsrud, A. J.

Number of pages: 5

Pages: 2491-2495

Publication date: 2020

Host publication information

Title of host publication: 42nd Annual International Conferences of the IEEE Engineering in Medicine and Biology Society : Enabling Innovative Technologies for Global Healthcare, EMBC 2020

Publisher: IEEE

ISBN (Print): 978-1-7281-1991-5

ISBN (Electronic): 9781728119908

Publication series

Name: Annual International Conference of the IEEE Engineering in Medicine and Biology Society

Volume: 2020-July

ISSN (Electronic): 1558-4615

ASJC Scopus subject areas: Signal Processing, Biomedical Engineering, Computer Vision and Pattern Recognition, Health Informatics

DOIs:

10.1109/EMBC44109.2020.9175306

Source: Scopus

Source ID: 85091006468

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Viewport-dependent delivery schemes for stereoscopic panoramic video

Stereoscopic panoramic or omnidirectional video is a key ingredient for an immersive experience in virtual reality applications. The user views only a portion of the omnidirectional scene at each time instant, hence streaming the whole stereoscopic panoramic or omnidirectional video in high quality is not necessary and will consume an unnecessary high bandwidth usage. In order to alleviate the problem of bandwidth wastage, viewport-dependent delivery schemes have been proposed, in which the part of the captured scene that is within the viewer's field of view is delivered at highest quality while the rest of the scene in lower quality. The low quality content is visible only after fast head movements for a short period, until the next periodic intra-coded picture that can be used for switching viewpoints is available. This paper proposes viewport-dependent delivery schemes for streaming of stereoscopic panoramic or omnidirectional video by using region-of-interest coding methods of MV-HEVC and SHVC standards. The proposed schemes avoid the need for frequent intra-coded pictures, and consequently in the performed experiments the streaming bitrate is reduced by more than 50% on average for the best schemes compared to a simulcast delivery method.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Nokia

Contributors: Ghaznavi-Youvalari, R., Hannuksela, M. M., Aminlou, A., Gabbouj, M.

Number of pages: 4

Pages: 1-4

Publication date: 2 Feb 2018

Host publication information

Title of host publication: 3DTV-CON 2017 - 3D True Vision v2 : Research and Applications in Future 3D Media

Publisher: IEEE

ISBN (Electronic): 9781538616352

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Networks and Communications, Computer Vision and Pattern Recognition, Human-Computer Interaction, Electrical and Electronic Engineering

Keywords: HEVC, MV-HEVC, panoramic video streaming, SHVC, video coding, Virtual reality

DOIs:

10.1109/3DTV.2017.8280404

Bibliographical note

EXT="Ghaznavi-Youvalari, Ramin"

EXT="Aminlou, Alireza"

Source: Scopus

Source ID: 85046375176

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

On the control of the KNTU CDRPM: A cable driven redundant parallel manipulator

This paper is devoted to the control of a cable driven redundant parallel manipulator, which is a challenging problem due to the optimal resolution of its inherent redundancy. Additionally to complicated forward kinematics, having a wide workspace makes it difficult to directly measure the pose of the end-effector. The goal of the controller is trajectory tracking in a large and singular free workspace, and to guarantee that the cables are always under tension. A control topology is proposed in this paper which is capable to fulfill the stringent positioning requirements for these type of manipulators. Closed-loop performance of various control topologies are compared by simulation of the closed-loop dynamics of the KNTU CDRPM, while the equations of parallel manipulator dynamics are implicit in structure and only special integration routines can be used for their integration. It is shown that the proposed joint space controller is capable to satisfy the required tracking performance, despite the inherent limitation of task space pose measurement.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Intelligent Hydraulics and Automation

Contributors: Gholami, P., Aref, M. M., Taghirad, H. D.

Number of pages: 6

Pages: 2404-2409

Publication date: 2008

Host publication information

Title of host publication: 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS

Article number: 4650740

ISBN (Print): 9781424420582

ASJC Scopus subject areas: Artificial Intelligence, Computer Vision and Pattern Recognition, Control and Systems Engineering, Electrical and Electronic Engineering

DOIs:

10.1109/IROS.2008.4650740

URLs:

<http://www.scopus.com/inward/record.url?scp=69549120718&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 69549120718

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Face typing: Vision-based perceptual interface for hands-free text entry with a scrollable virtual keyboard

We present a novel vision-based perceptual user interface for hands-free text entry that utilizes face detection and visual gesture detection to manipulate a scrollable virtual keyboard. A thorough experimentation was undertaken to quantitatively define a performance of the interface in hands-free pointing, selection and scrolling tasks. The experiments were conducted with nine participants in laboratory conditions. Several face and head gestures were examined for detection robustness and user convenience. The system gave a reasonable performance in terms of high gesture detection rate and small false alarm rate. The participants reported that a new interface was easy to understand and operate. Encouraged by these results, we discuss advantages and constraints of the interface and suggest possibilities for design improvements.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA)

Contributors: Gizatdinova, Y., Spakov, Ō., Surakka, V.

Number of pages: 7

Pages: 81-87

Publication date: 2012

Host publication information

Title of host publication: 2012 IEEE Workshop on the Applications of Computer Vision, WACV 2012

Article number: 6162997

ISBN (Print): 9781467302333

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Science Applications

DOIs:

10.1109/WACV.2012.6162997

URLs:

<http://www.scopus.com/inward/record.url?scp=84860699077&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84860699077

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Facial expression classification based on local spatiotemporal edge and texture descriptors

Facial expressions are emotionally, socially and otherwise meaningful reflective signals in the face. Facial expressions play a critical role in human life, providing an important channel of nonverbal communication. Automation of the entire process of expression analysis can potentially facilitate human-computer interaction, making it to resemble mechanisms of human-human communication. In this paper, we present an ongoing research that aims at development of a novel spatiotemporal approach to expression classification in video. The novelty comes from a new facial representation that is based on local spatiotemporal feature descriptors. In particular, a combined dynamic edge and texture information is used for reliable description of both appearance and motion of the expression. Support vector machines are utilized to perform a final expression classification. The planned experiments will further systematically evaluate the performance of the developed method with several databases of complex facial expressions.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA), Univ of Oulu

Contributors: Gizatdinova, Y., Surakka, V., Zhao, G., Mäkinen, E., Raisamo, R.

Publication date: 2011

Host publication information

Title of host publication: Selected Papers from the Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research - Digital Edition, MB'10

Article number: 21

ISBN (Print): 9781605589268

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Action unit, Emotion, Expression classification, Facial expression, Human behaviour understanding, Local binary pattern, Local oriented edge, Spatiotemporal descriptor

DOIs:

10.1145/1931344.1931365

URLs:

<http://www.scopus.com/inward/record.url?scp=79952499491&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 79952499491

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Reducing the number of unit tests with design by contract

Design by Contract (DbC) and unit testing (UT) are complementary approaches to improve the belief of correctness and the quality of the software. The interplay between the two techniques has been studied previously, e.g., in the use of test oracles and test automation. However, we propose that DbC should drive the UT to become more cost-effective. The paper demonstrates some means for this approach by showing how to test a mapping data structure entirely with just one unit test script.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Managing digital industrial transformation (mDIT), Department of Information Technology, Turku Centre for Computer Science, University of Turku

Contributors: Hakonen, H., Hyrnsalmi, S., Järvi, A.

Number of pages: 6

Pages: 161-166

Publication date: 2011

Host publication information

Title of host publication: Computer Systems and Technologies - 12th International Conference, CompSysTech'11 - Proceedings

Volume: 578

ISBN (Print): 9781450309172

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: contract cohesion graph, design by contract, unit testing

DOIs:

10.1145/2023607.2023635

Source: Scopus

Source ID: 80052810613

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Bayesian receiver operating characteristic metric for linear classifiers

We propose a novel classifier accuracy metric: the Bayesian Area Under the Receiver Operating Characteristic Curve (CBAUC). The method estimates the area under the ROC curve and is related to the recently proposed Bayesian Error Estimator. The metric can assess the quality of a classifier using only the training dataset without the need for computationally expensive cross-validation. We derive a closed-form solution of the proposed accuracy metric for any linear binary classifier under the Gaussianity assumption, and study the accuracy of the proposed estimator using simulated and real-world data. These experiments confirm that the closed-form CBAUC is both faster and more accurate than conventional AUC estimators.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Research group: Computational Systems Biology, Computing Sciences, AI Virtanen Institute for Molecular Sciences, University of Eastern Finland

Contributors: Hassan, S. S., Huttunen, H., Niemi, J., Tohka, J.

Number of pages: 8

Pages: 52-59

Publication date: 1 Dec 2019

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters

Volume: 128

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2019): CiteScore 6.3 SJR 0.848 SNIP 2.021

Original language: English

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Artificial Intelligence

Keywords: Bayesian error estimation, Classification, Receiver operating characteristic curve

DOIs:

10.1016/j.patrec.2019.07.016

Bibliographical note

EXT="Tohka, Jussi"

Source: Scopus

Source ID: 85071016385

Research output: Contribution to journal > Article > Scientific > peer-review

Simple gaze gestures and the closure of the eyes as an interaction technique

We created a set of gaze gestures that utilize the following three elements: simple one-segment gestures, off-screen space, and the closure of the eyes. These gestures are to be used as the moving tool in a gaze-only controlled drawing application. We tested our gaze gestures with 24 participants and analyzed the gesture durations, the accuracy of the stops, and the gesture performance. We found that the difference in gesture durations between short and long gestures was so small that there is no need to choose between them. The stops made by closing both eyes were accurate, and the input method worked well for this purpose. With some adjustments and with the possibility for personal settings, the gesture performance and the accuracy of the stops can become even better.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA)

Contributors: Heikkilä, H., Rähä, K. J.

Number of pages: 8

Pages: 147-154

Publication date: 2012

Host publication information

Title of host publication: Proceedings - ETRA 2012: Eye Tracking Research and Applications Symposium
ISBN (Print): 9781450312257
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems
Keywords: closure of both eyes, eye tracking, gaze control, gaze gestures, gaze-based interaction, off-screen space
DOIs:
10.1145/2168556.2168579
URLs:
<http://www.scopus.com/inward/record.url?scp=84862701036&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84862701036
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Learnings from the Finnish game industry

The motivation behind our research was the rapid growth and business wins of world-class Finnish game companies, like Supercell, as well as the success of other game companies in Finland. In particular, Supercell's growth is something that has not been heard of before and this raised the interest to research what game companies have been doing right. Supercell is not the only Finnish success. Rovio is also well known and has the roots for success from few years before. There are also other game companies in Finland that have succeeded and this motivated us to investigate what is happening behind the game industry and what could be learned from there that could be applied to other software industry as well. In order to explore and explain the different success factors, we interviewed the following eight Finnish game companies: Rovio Entertainment, Fingersoft, TicBits, Boomlagoon, 10tons, Tribeflame, Star Arcade and Mountain Sheep. In addition, we investigated public sources, like interviews given to newspapers and books written about companies. These sources cover well Supercell as they have given numerous public interviews to journalists. Similarly, Remedy was analysed based on public sources. Based on the results we recognised some 30 patterns that, depending on the context, could be used in other organisations as well. The patterns include the applicable context where they can be used, driving forces (and counterforces) that should be recognised, the problem they are solving and the solution to the problem coupled with the key enablers. Furthermore, narrative stories based on the interviews and public sources are included.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Pervasive Computing, Research area: Information security, University of Helsinki, Jyväskylän yliopisto
Contributors: Helenius, M., Kettunen, P., Frank, L.
Publication date: 7 Apr 2016

Host publication information

Title of host publication: Proceedings of the 10th Travelling Conference on Pattern Languages of Programs, VikingPLOP 2016
Publisher: ACM
Article number: a12
ISBN (Electronic): 9781450342001
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
DOIs:
10.1145/3022636.3022648

Bibliographical note

EXT="Frank, Lauri"
Source: Scopus
Source ID: 85015616047
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Sparse modelling and predictive coding of subaperture images for lossless plenoptic image compression

This paper studies the lossless compression of rectified light-field images captured by plenoptic cameras, exploiting the high similarity existing between the subaperture images, or views, composing the light-field image. The encoding is predictive, where one sparse predictor is designed for every region of a view, using as regressors the pixels from the already transmitted views. As a first step, consistent segmentations for all subaperture images are constructed, defining the regions as connected components in the quantized depth map of the central view, and then propagating them to all side views. The sparse predictors are able to take into account the small horizontal and vertical disparities between regions in corresponding close-by views and perform optimal least squares interpolation accounting implicitly for fractional disparities. The optimal structure of the sparse predictor is selected for each region based on an implementable description length. The encoding of the views is done sequentially starting from the central view and the scheme produces results better than standard lossless compression methods utilized directly on the full lightfield image or applied to the views in a similar sequential order as our method.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: Algebraic and Algorithmic Methods in Signal Processing AAMSP, Research group: Signal Interpretation and Compression-SIC, Signal Processing Research Community (SPRC), University of California San Diego

Contributors: Helin, P., Astola, P., Rao, B., Tabus, I.

Publication date: 22 Aug 2016

Host publication information

Title of host publication: 2016 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video, 3DTV-CON 2016

Publisher: IEEE COMPUTER SOCIETY PRESS

ISBN (Electronic): 9781509033133

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Networks and Communications, Computer Vision and Pattern Recognition, Human-Computer Interaction, Electrical and Electronic Engineering

Keywords: depth map warping, light-field coding, lossless compression, plenoptics, sparse prediction

DOIs:

10.1109/3DTV.2016.7548953

Source: Scopus

Source ID: 84987803027

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Recording and analyzing in-browser programming sessions

In this paper, we report on the analysis of a novel type of automatically recorded detailed programming session data collected on a university-level web programming course. We present a method and an implementation of collecting rich data on how students learning to program edit and execute code and explore its use in examining learners' behavior. The data collection instrument is an in-browser Python programming environment that integrates an editor, an execution environment, and an interactive Python console and is used to deliver programming assignments with automatic feedback. Most importantly, the environment records learners' interaction within it. We have implemented tools for viewing these traces and demonstrate their potential in learning about the programming processes of learners and of benefiting computing education research and the teaching of programming.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Regulation of learning and active learning methods (REALMEE), Aalto University, Department of Computer Science and Eng.

Contributors: Helminen, J., Ihantola, P., Karavirta, V.

Number of pages: 10

Pages: 13-22

Publication date: 2013

Host publication information

Title of host publication: Proceedings - 13th Koli Calling International Conference on Computing Education Research, Koli Calling 2013

ISBN (Print): 9781450324823

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: computer science education, computing education research, programming assignment, programming session, Python, web based programming environment

DOIs:

10.1145/2526968.2526970

URLs:

<http://www.scopus.com/inward/record.url?scp=84889581968&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84889581968

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Automatic image-based detection and inspection of paper fibres for grasping

An automatic computer vision algorithm that detects individual paper fibres from an image, assesses the possibility of grasping the detected fibres with microgrippers and detects the suitable grasping points is presented. The goal of the algorithm is to enable automatic fibre manipulation for mechanical characterisation, which has traditionally been slow manual work. The algorithm classifies the objects in images based on their morphology, and detects the proper grasp points from the individual fibres by applying given geometrical constraints. The authors test the ability of the algorithm to

detect the individual fibres with 35 images containing more than 500 fibres in total, and also compare the graspability analysis and the calculated grasp points with the results of an experienced human operator with 15 images that contain a total of almost 200 fibres. The detection results are outstanding, with fewer than 1% of fibres missed. The graspability analysis gives sensitivity of 0.83 and specificity of 0.92, and the average distance between the grasp points of the human and the algorithm is 220 µm. Also, the choices made by the algorithm are much more consistent than the human choices.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Automation Science and Engineering, Integrated Technologies for Tissue Engineering Research (ITTE)

Contributors: Hirvonen, J., Kallio, P.

Number of pages: 7

Pages: 588-594

Publication date: 1 Aug 2015

Peer-reviewed: Yes

Publication information

Journal: IET Computer Vision

Volume: 9

Issue number: 4

ISSN (Print): 1751-9632

Ratings:

Scopus rating (2015): CiteScore 2.3 SJR 0.3 SNIP 1.218

Original language: English

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Software

Electronic versions:

IET_post_print

DOIs:

10.1049/iet-cvi.2014.0416

URLs:

<http://urn.fi/URN:NBN:fi:tyy-201603243758>

Source: Scopus

Source ID: 84938530267

Research output: Contribution to journal > Article > Scientific > peer-review

Three patterns for user involvement in startups

Creating products in software startups consists of a great deal of uncertainty combined with little resources. Rapid validation of created solutions with the potential customers is essential to startups. However, often startups lack people with skills needed for the validation. We present three patterns that help in involving users to gain meaningful feedback and learning. First, the feedback has to be gotten from the right people and the right questions have to be asked. Furthermore, if the feedback is collected with a prototype, often called a Minimum Viable Product, users should be able to give feedback of the actual idea, not to any roughness caused by the immaturity and the prototypishness of the product.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Pervasive Computing, Research area: Software engineering

Contributors: Hokkanen, L., Leppänen, M.

Publication date: 8 Jul 2015

Host publication information

Title of host publication: Proceedings of the 20th European Conference on Pattern Languages of Programs, EuroPLoP 2015

Publisher: ACM

Article number: a51

ISBN (Electronic): 9781450338479

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Lean, Startups, User experience

DOIs:

10.1145/2855321.2855373

Source: Scopus

Source ID: 84982794686

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

User Experience Study of 360° Music Videos on Computer Monitor and Virtual Reality Goggles

360° videos are increasingly used for media and entertainment, but the best practices for editing them are not yet well established. In this paper, we present a study in which we investigated the user experience of 360° music videos viewed on computer monitor and VR goggles. The research was conducted in the form of a laboratory experiment with 20 test participants. During the within-subject study, participants watched and evaluated four versions of the same 360° music video with a different cutting rate. Based on the results, an average cutting rate of 26 seconds delivered the highest-quality user experience both for computer monitor and VR goggles. The cutting rate matched with participants' mental models, and there was enough time to explore the environment without getting bored. Faster cutting rates made the users nervous, and a video consisting of a single shot was considered to be too static and boring.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Tampere University, Tampere University of Applied Sciences

Contributors: Holm, J., Väänänen, K., Remans, M. M. R.

Number of pages: 7

Pages: 81-87

Publication date: 1 Jul 2019

Host publication information

Title of host publication: Information Visualization - Biomedical Visualization and Geometric Modelling and Imaging, IV 2019

Publisher: IEEE

Editors: Banissi, E., Ursyn, A., McK. Bannatyne, M. W., Datia, N., Pires, J. M., Francese, R., Sarfraz, M., Wyeld, T. G., Bouali, F., Venturin, G., Azzag, H., Lebbah, M., Trutschl, M., Cvek, U., Muller, H., Nakayama, M., Kernbach, S., Caruccio, L., Risi, M., Erra, U., Vitiello, A., Rossano, V.

ISBN (Electronic): 9781728128382

Publication series

Name: Proceedings of the International Conference on Information Visualisation

ISSN (Print): 1093-9547

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition

Keywords: 360° video, cutting rate, hmd, music video, virtual reality, virtual reality goggles

DOIs:

10.1109/IV.2019.00023

Bibliographical note

jufoid=58079

EXT="Holm, Jukka"

INT=comp,"Remans, Mohammad Mushfiqur Rahman"

Source: Scopus

Source ID: 85072286445

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

User experience of stereo and spatial audio in 360° live music videos

360° music videos are becoming prevalent in music entertainment. Still, academic studies of the 360° live music experience covering both audio and visual experience are scarce. In this paper, we present a study of user experience of stereo and spatial audio in a 360° live music video setting with two different display types. The research was conducted in the form of a laboratory experiment, in which 20 participants watched and evaluated stereo and spatial audio versions of the same music video using a flat computer display and a head-mounted display (HMD). Based on the results, spatial audio combined with HMD scored highest in the quantitative metrics of perceived audio quality, presence, and overall listening experience. However, qualitative findings reveal that this combination does not fit well with users' listening habits. While nine participants preferred to use headphones to listen to music, thirteen participants viewed music listening as a secondary task-making the use of HMDs less suitable.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Tampere University, Tampere University of Applied Sciences

Contributors: Holm, J., Väänänen, K., Battah, A.

Number of pages: 8

Pages: 134-141

Publication date: 2020

Host publication information

Title of host publication: AcademicMindtrek 2020 - Proceedings of the 23rd International Academic Mindtrek Conference : January 2020, Tampere

Publisher: ACM

ISBN (Electronic): 9781450377744

ASJC Scopus subject areas: Software, Human-Computer Interaction, Computer Vision and Pattern Recognition, Computer Networks and Communications

Keywords: 360° video, ambisonics, head-mounted display, music video, spatial audio, stereo, user experience, virtual reality

DOIs:

10.1145/3377290.3377291

Bibliographical note

EXT="Holm, Jukka"

INT=comp,"Battah, Anas"

Source: Scopus

Source ID: 85080964162

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Social and privacy aspects of a system for collaborative public expression

In this paper, we are concerned with how a real-world social situation shapes the interaction with a novel technology that combines collocated mobile phone and public display use for groups of people. We present a user study of a system that allows collaborative creation and sharing of comic strips on public displays in a social setting such as a pub or café. The system utilizes mobile phones and public displays for shared collaborative expression between collocated users. A user study spanning three sessions was conducted in real-world settings: one during the social event following a seminar on games research and two in a bar on a regular weekday evening. We present and discuss our findings with respect to how the larger social situation and location influenced the interaction with the system, the collaboration between participants of a team, how people moved between different roles (i.e., actor, spectator and bystander), and the privacy issues it evoked from participants.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Mathematical modelling with wide societal impact (MathImpact), Nokia

Contributors: Holopainen, J., Lucero, A., Saarenpää, H., Nummenmaa, T., Ali, A. E., Jokela, T.

Publication date: 2011

Host publication information

Title of host publication: Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, ACE 2011

Article number: 23

ISBN (Print): 9781450308274

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Collaborative interaction, Evaluation, Mobile phones, Public interfaces, Social context

DOIs:

10.1145/2071423.2071452

URLs:

<http://www.scopus.com/inward/record.url?scp=84855410287&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84855410287

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Asymmetric luminance based filtering for stereoscopic video compression

Asymmetric stereo video coding is a well-known enhancement technique for efficient 3D video rate scaling, taking advantage from the binocular suppression theory. Usually in asymmetric video coding one view is encoded in a higher spatial/ temporal resolution or quality, while the auxiliary view is encoded with a lower resolution or quality. In this paper a novel asymmetric video coding approach is proposed to enhance the stereoscopic video compression efficiency. A regionally adaptive smoothing filter is applied to dark pixels of one view while the same filter is only applied to the light pixels of the other view. The location and the strength of the smoothing filters are determined according to the texture characteristics and the degree of the brightness of each individual pixel within the image. A series of systematic subjective tests were conducted, confirming that no quality degradation is perceptible by application of such filters. This is while the objective measurements show a Bjontegaard delta bitrate reduction of up to 26.6% and with an average of 16.8%.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: Video, Research Community on Data-to-Decision (D2D), Nokia

Contributors: Homayouni, M., Aflaki, P., Hannuksela, M. M., Gabbouj, M.

Publication date: 5 Feb 2015

Host publication information

Title of host publication: 2014 International Conference on 3D Imaging (IC3D)

Publisher: IEEE

ISBN (Print): 9781479980239

ASJC Scopus subject areas: Organizational Behavior and Human Resource Management, Computer Vision and Pattern Recognition

Keywords: Asymmetric, low-pass filter, MVC, Stereoscopic video, subjective quality assessment

DOIs:

10.1109/IC3D.2014.7032577

Source: Scopus

Source ID: 84930446866

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

A survey on aims and environments of diversification and obfuscation in software security

Diversification and obfuscation methods are promising approaches used to secure software and prevent malware from functioning. Diversification makes each software instance unique so that malware attacks cannot rely on the knowledge of the program's execution environment and/or internal structure anymore. We present a systematic literature review on the state-of-the-art of diversification and obfuscation research aiming to improve software security between 1993 and 2014. As the result of systematic search, in the final phase, 209 related papers were included in this study. In this study we focus on two specific research questions: what are the aims of diversification and obfuscation techniques and what are the environments they are applied to. The former question includes the languages and the execution environments that can benefit from these two techniques, while the second question presents the goals of the techniques and also the type of attacks they mitigate. is held by the owner/author(s). Publication rights licensed to ACM.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Turun Yliopisto/Turun Biomateriaalikeskus

Contributors: Hosseinzadeh, S., Rauti, S., Laurén, S., Mäkelä, J. M., Holvitie, J., Hyrynsalmi, S., Leppänen, V.

Number of pages: 8

Pages: 113-120

Publication date: 23 Jun 2016

Host publication information

Title of host publication: Computer Systems and Technologies 17th International Conference, CompSysTech 2016 - Proceedings

Volume: 1164

Publisher: Association for Computing Machinery

ISBN (Electronic): 9781450341820

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Diversification, Obfuscation, Software security, Systematic literature review (SLR)

DOIs:

10.1145/2983468.2983479

URLs:

<http://www.scopus.com/inward/record.url?scp=85000983786&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 85000983786

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Lossless compression of regions-of-interest from retinal images

This paper presents a lossless compression method performing separately the compression of the vessels and of the remaining part of eye fundus in retinal images. Retinal images contain valuable information sources for several distinct medical diagnosis tasks, where the features of interest can be e.g. the cotton wool spots in the eye fundus, or the volume of the vessels over concentric circular regions. It is assumed that one of the existent segmentation methods provided the segmentation of the vessels. The proposed compression method transmits losslessly the segmentation image, and then transmits the eye fundus part, or the vessels image, or both, conditional on the vessels segmentation. The independent

compression of the two color image segments is performed using a sparse predictive method. Experiments are provided over a database of retinal images containing manual and estimated segmentations. The codelength of encoding the overall image, including the segmentation and the image segments, proves to be better than the codelength for the entire image obtained by JPEG2000 and other publicly available compressors.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: Signal Interpretation and Compression-SIC, Signal Processing Research Community (SPRC)

Contributors: Hukkanen, J., Astola, P., Tabus, I.

Publication date: 22 Jan 2015

Host publication information

Title of host publication: EUVIP 2014 - 5th European Workshop on Visual Information Processing

Publisher: The Institute of Electrical and Electronics Engineers, Inc.

ISBN (Print): 9781479945726

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Information Systems, Signal Processing

Keywords: lossless compression, region of interest, retinal images, sparse prediction

DOIs:

10.1109/EUVIP.2014.7018394

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Model selection for linear classifiers using Bayesian error estimation

Regularized linear models are important classification methods for high dimensional problems, where regularized linear classifiers are often preferred due to their ability to avoid overfitting. The degree of freedom of the model is determined by a regularization parameter, which is typically selected using counting based approaches, such as K-fold cross-validation. For large data, this can be very time consuming, and, for small sample sizes, the accuracy of the model selection is limited by the large variance of CV error estimates. In this paper, we study the applicability of a recently proposed Bayesian error estimator for the selection of the best model along the regularization path. We also propose an extension of the estimator that allows model selection in multiclass cases and study its efficiency with L-1 regularized logistic regression and L-2 regularized linear support vector machine. The model selection by the new Bayesian error estimator is experimentally shown to improve the classification accuracy, especially in small sample-size situations, and is able to avoid the excess variability inherent to traditional cross-validation approaches. Moreover, the method has significantly smaller computational complexity than cross-validation. (C) 2015 Elsevier Ltd. All rights reserved.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research group: Vision, Research Community on Data-to-Decision (D2D), Universidad Carlos III de Madrid

Contributors: Huttunen, H., Tohka, J.

Number of pages: 10

Pages: 3739-3748

Publication date: Nov 2015

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition

Volume: 48

Issue number: 11

ISSN (Print): 0031-3203

Ratings:

Scopus rating (2015): CiteScore 8.6 SJR 1.579 SNIP 2.996

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Logistic regression, Support vector machine, Regularization, Bayesian error estimator, Linear classifier, MULTINOMIAL LOGISTIC-REGRESSION, SUPPORT VECTOR MACHINES, CLASSIFICATION, PERFORMANCE, BOUNDS

DOIs:

10.1016/j.patcog.2015.05.005

Bibliographical note

EXT="Tohka, Jussi"

Source: Scopus

Source ID: 84937812363

Research output: Contribution to journal › Article › Scientific › peer-review

Gaze gestures or dwell-based interaction?

The two cardinal problems recognized with gaze-based interaction techniques are: how to avoid unintentional commands, and how to overcome the limited accuracy of eye tracking. Gaze gestures are a relatively new technique for giving commands, which has the potential to overcome these problems. We present a study that compares gaze gestures with dwell selection as an interaction technique. The study involved 12 participants and was performed in the context of using an actual application. The participants gave commands to a 3D immersive game using gaze gestures and dwell icons. We found that gaze gestures are not only a feasible means of issuing commands in the course of game play, but they also exhibited performance that was at least as good as or better than dwell selections. The gesture condition produced less than half of the errors when compared with the dwell condition. The study shows that gestures provide a robust alternative to dwell-based interaction with the reliance on positional accuracy being substantially reduced.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA), De Montfort University

Contributors: Hyrskykari, A., Istance, H., Vickers, S.

Number of pages: 4

Pages: 229-232

Publication date: 2012

Host publication information

Title of host publication: Proceedings - ETRA 2012: Eye Tracking Research and Applications Symposium

ISBN (Print): 9781450312257

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems

Keywords: assistive input devices, eye tracking, gaze and gaming, gaze gestures, physically disabled user groups

DOIs:

10.1145/2168556.2168602

URLs:

<http://www.scopus.com/inward/record.url?scp=84862671730&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84862671730

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Multi-modal dense video captioning

Dense video captioning is a task of localizing interesting events from an untrimmed video and producing textual description (captions) for each localized event. Most of the previous works in dense video captioning are solely based on visual information and completely ignore the audio track. However, audio, and speech, in particular, are vital cues for a human observer in understanding an environment. In this paper, we present a new dense video captioning approach that is able to utilize any number of modalities for event description. Specifically, we show how audio and speech modalities may improve a dense video captioning model. We apply automatic speech recognition (ASR) system to obtain a temporally aligned textual description of the speech (similar to subtitles) and treat it as a separate input alongside video frames and the corresponding audio track. We formulate the captioning task as a machine translation problem and utilize recently proposed Transformer architecture to convert multi-modal input data into textual descriptions. We demonstrate the performance of our model on ActivityNet Captions dataset. The ablation studies indicate a considerable contribution from audio and speech components suggesting that these modalities contain substantial complementary information to video frames. Furthermore, we provide an in-depth analysis of the ActivityNet Caption results by leveraging the category tags obtained from original YouTube videos. Code is publicly available: github.com/v-iashin/MDVC.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Research group: Artificial Intelligence and Vision - AIV

Contributors: Iashin, V., Rahtu, E.

Number of pages: 10

Pages: 4117-4126

Publication date: 2020

Host publication information

Title of host publication: Proceedings - 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2020

Publisher: IEEE
ISBN (Print): 978-1-7281-9361-8
ISBN (Electronic): 9781728193601

Publication series

Name: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops
ISSN (Print): 2160-7508
ISSN (Electronic): 2160-7516
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Electrical and Electronic Engineering
DOIs:
10.1109/CVPRW50498.2020.00487

Bibliographical note

JUFID=70566
Source: Scopus
Source ID: 85090152305
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

How to study programming on mobile touch devices - Interactive Python code exercises

Scaffolded learning tasks where programs are constructed from predefined code fragments by dragging and dropping them (i.e. Parsons problems) are well suited to mobile touch devices, but quite limited in their applicability. They do not adequately cater for different approaches to constructing a program. After studying solutions to automatically assessed programming exercises, we found out that many different solutions are composed of a relatively small set of mutually similar code lines. Thus, they can be constructed by using the drag-and-drop approach if only it was possible to edit some small parts of the predefined fragments. Based on this, we have designed and implemented a new exercise type for mobile devices that builds on Parsons problems and falls somewhere between their strict scaffolding and full-blown coding exercises. In these exercises, we can gradually fade the scaffolding and allow programs to be constructed more freely so as not to restrict thinking and limit creativity too much while still making sure we are able to deploy them to small-screen mobile devices. In addition to the new concept and the related implementation, we discuss other possibilities of how programming could be practiced on mobile devices.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Regulation of learning and active learning methods (REALMEE), Department of Computer Science and Eng., Aalto University
Contributors: Ihantola, P., Helminen, J., Karavirta, V.
Number of pages: 8
Pages: 51-58
Publication date: 2013

Host publication information

Title of host publication: Proceedings - 13th Koli Calling International Conference on Computing Education Research, Koli Calling 2013
ISBN (Print): 9781450324823
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: learning, mLearning, mobile learning, mobile touch devices, parsons problem, parsons puzzle, programming, Python, teaching
DOIs:
10.1145/2526968.2526974
URLs:
<http://www.scopus.com/inward/record.url?scp=84889570829&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84889570829
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

On the kernel Extreme Learning Machine classifier

In this paper, we discuss the connection of the kernel versions of the ELM classifier with infinite Single-hidden Layer Feedforward Neural networks and show that the original ELM kernel definition can be adopted for the calculation of the ELM kernel matrix for two of the most common activation functions, i.e., the RBF and the sigmoid functions. In addition, we show that a low-rank decomposition of the kernel matrix defined on the input training data can be exploited in order to determine an appropriate ELM space for input data mapping. The ELM space determined from this process can be subsequently used for network training using the original ELM formulation. Experimental results denote that the adoption of the low-rank decomposition-based ELM space determination leads to enhanced performance, when compared to the

standard choice, i.e., random input weights generation.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics

Contributors: Iosifidis, A., Tefas, A., Pitas, I.

Number of pages: 7

Pages: 11-17

Publication date: 1 Mar 2015

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters

Volume: 54

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2015): CiteScore 5.1 SJR 0.95 SNIP 2.002

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Extreme learning machine, Infinite networks, Single-hidden layer networks

DOIs:

10.1016/j.patrec.2014.12.003

URLs:

<http://www.scopus.com/inward/record.url?scp=84920068822&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84920068822

Research output: Contribution to journal › Article › Scientific › peer-review

The MOBISERV-AIIA eating and drinking multi-view database for vision-based assisted living

Assisted living has a particular social importance in most developed societies, due to the increased life expectancy of the general population and the ensuing ageing problems. It has also importance for the provision of improved home care in cases of disabled persons or persons suffering from certain diseases that have high social impact. In this context, the development of computer vision systems capable to identify human eating and drinking activity can be really useful in order to prevent undernourishment/malnutrition and dehydration in a smart home environment targeting to extend independent living of older persons in the early stage of dementia. In this paper, we first describe the human centered interface specifications and implementations for such a system, which can be supported by ambient intelligence and robotic technologies. We, subsequently, describe a multi-view eating and drinking activity recognition database that has been created in order to facilitate research towards this direction. The database has been created by using four cameras in order to produce multi-view videos, each depicting one of twelve persons having a meal, resulting to a database size equal to 59.68 hours in total. Various types of meals have been recorded, i.e., breakfast, lunch and fast food. Moreover, the persons have different sizes, clothing and are of different sex. The database has been annotated in a frame base in terms of person ID and activity class. We hope that such a database will serve as a benchmark data set for computer vision researchers in order to devise methods targeting to this important application.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Research Community on Data-to-Decision (D2D), School of Dentistry, Aristotle University of Thessaloniki, Department of Informatics

Contributors: Iosifidis, A., Marami, E., Tefas, A., Pitas, I., Lyroudia, K.

Number of pages: 20

Pages: 254-273

Publication date: 1 Mar 2015

Peer-reviewed: Yes

Publication information

Journal: Journal of Information Hiding and Multimedia Signal Processing

Volume: 6

Issue number: 2

ISSN (Print): 2073-4212

Ratings:

Scopus rating (2015): CiteScore 2.6 SJR 0.414 SNIP 1.483

Original language: English
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Software
Keywords: Activity recognition, Multiview video database, Nutrition assistance, Smart home environment
Source: Scopus
Source ID: 84911457722
Research output: Contribution to journal › Article › Scientific › peer-review

Kernel reference discriminant analysis

Linear Discriminant Analysis (LDA) and its nonlinear version Kernel Discriminant Analysis (KDA) are well-known and widely used techniques for supervised feature extraction and dimensionality reduction. They determine an optimal discriminant space for (non)linear data projection based on certain assumptions, e.g. on using normal distributions (either on the input or in the kernel space) for each class and employing class representation by the corresponding class mean vectors. However, there might be other vectors that can be used for classes representation, in order to increase class discrimination in the resulted feature space. In this paper, we propose an optimization scheme aiming at the optimal class representation, in terms of Fisher ratio maximization, for nonlinear data projection. Compared to the standard approach, the proposed optimization scheme increases class discrimination in the reduced-dimensionality feature space and achieves higher classification rates in publicly available data sets.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics
Contributors: Iosifidis, A., Tefas, A., Pitas, I.
Number of pages: 7
Pages: 85-91
Publication date: 1 Nov 2014
Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters
Volume: 49
ISSN (Print): 0167-8655
Ratings:
Scopus rating (2014): CiteScore 4.3 SJR 0.73 SNIP 2.131
Original language: English
ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing
Keywords: Kernel Discriminant Analysis, Kernel Spectral Regression, Optimized class representation
DOIs:
10.1016/j.patrec.2014.06.013
Source: Scopus
Source ID: 84904957982
Research output: Contribution to journal › Article › Scientific › peer-review

Discriminant Bag of Words based representation for human action recognition

In this paper we propose a novel framework for human action recognition based on Bag of Words (BoWs) action representation, that unifies discriminative codebook generation and discriminant subspace learning. The proposed framework is able to, naturally, incorporate several (linear or non-linear) discrimination criteria for discriminant BoWs-based action representation. An iterative optimization scheme is proposed for sequential discriminant BoWs-based action representation and codebook adaptation based on action discrimination in a reduced dimensionality feature space where action classes are better discriminated. Experiments on five publicly available data sets aiming at different application scenarios demonstrate that the proposed unified approach increases the codebook discriminative ability providing enhanced action classification performance.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics
Contributors: Iosifidis, A., Tefas, A., Pitas, I.
Number of pages: 8
Pages: 185-192
Publication date: 1 Nov 2014
Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters

Volume: 49

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2014): CiteScore 4.3 SJR 0.73 SNIP 2.131

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Bag of Words, Codebook learning, Discriminant learning

DOIs:

10.1016/j.patrec.2014.07.011

Source: Scopus

Source ID: 84907347636

Research output: Contribution to journal › Article › Scientific › peer-review

Semi-supervised classification of human actions based on neural networks

In this paper, we propose a novel algorithm for Single-hidden Layer Feed forward Neural networks training which is able to exploit information coming from both labeled and unlabeled data for semi-supervised action classification. We extend the Extreme Learning Machine algorithm by incorporating appropriate regularization terms describing geometric properties and discrimination criteria of the training data representation in the ELM space to this end. The proposed algorithm is evaluated on human action recognition, where its performance is compared with that of other (semi-)supervised classification schemes. Experimental results on two publicly available action recognition databases denote its effectiveness.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics

Contributors: Iosifidis, A., Tefas, A., Pitas, I.

Number of pages: 6

Pages: 1336-1341

Publication date: 4 Dec 2014

Host publication information

Title of host publication: Proceedings - International Conference on Pattern Recognition

Publisher: The Institute of Electrical and Electronics Engineers, Inc.

ISBN (Print): 9781479952083

ASJC Scopus subject areas: Computer Vision and Pattern Recognition

DOIs:

10.1109/ICPR.2014.239

Source: Scopus

Source ID: 84919935864

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Active classification for human action recognition

In this paper, we propose a novel classification method involving two processing steps. Given a test sample, the training data residing to its neighborhood are determined. Classification is performed by a Single-hidden Layer Feedforward Neural network exploiting labeling information of the training data appearing in the test sample neighborhood and using the rest training data as unlabeled. By following this approach, the proposed classification method focuses the classification problem on the training data that are more similar to the test sample under consideration and exploits information concerning to the training set structure. Compared to both static classification exploiting all the available training data and dynamic classification involving data selection for classification, the proposed active classification method provides enhanced classification performance in two publicly available action recognition databases.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics

Contributors: Iosifidis, A., Tefas, A., Pitas, I.

Number of pages: 5

Pages: 3249-3253

Publication date: 2013

Host publication information

Title of host publication: 2013 IEEE International Conference on Image Processing, ICIP 2013 - Proceedings

Article number: 6738669

ISBN (Print): 9781479923410

ASJC Scopus subject areas: Computer Vision and Pattern Recognition

Keywords: Active classification, dynamic classification, Extreme Learning Machine, human action recognition, Single-hidden Layer Feedforward Neural network

DOIs:

10.1109/ICIP.2013.6738669

Source: Scopus

Source ID: 84897811368

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

View-independent human action recognition based on multi-view action images and discriminant learning

In this paper a novel view-independent human action recognition method is proposed. A multi-camera setup is used to capture the human body from different viewing angles. Actions are described by a novel action representation, the so-called multi-view action image (MVAI), which effectively addresses the camera viewpoint identification problem, i.e., the identification of the position of each camera with respect to the person's body. Linear Discriminant Analysis is applied on the MVAIs in order to map actions to a discriminant feature space where actions are classified by using a simple nearest class centroid classification scheme. Experimental results denote the effectiveness of the proposed action recognition approach.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics

Contributors: Iosifidis, A., Tefas, A., Pitas, I.

Publication date: 2013

Host publication information

Title of host publication: 2013 IEEE 11th IVMSP Workshop: 3D Image/Video Technologies and Applications, IVMSP 2013 - Proceedings

ISBN (Print): 9781467358583

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Computer Science Applications

Keywords: Discriminant Learning, Human Action Recognition, Multi-camera Setup, Multi-view Action Images

DOIs:

10.1109/IVMSPW.2013.6611931

Source: Scopus

Source ID: 84888154998

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Multi-view action recognition based on action volumes, fuzzy distances and cluster discriminant analysis

In this paper, we present a view-independent action recognition method exploiting a low computational-cost volumetric action representation. Binary images depicting the human body during action execution are accumulated in order to produce the so-called action volumes. A novel time-invariant action representation is obtained by exploiting the circular shift invariance property of the magnitudes of the Discrete Fourier Transform coefficients. The similarity of an action volume with representative action volumes is exploited in order to map it to a lower-dimensional feature space that preserves the action class properties. Discriminant learning is, subsequently, employed for further dimensionality reduction and action class discrimination. By using such an action representation, the proposed approach performs fast action recognition. By combining action recognition results coming from different view angles, high recognition rates are obtained. The proposed method is extended to interaction recognition, i.e., to human action recognition involving two persons. The proposed approach is evaluated on a publicly available action recognition database using experimental settings simulating situations that may appear in real-life applications, as well as on a new nutrition support action recognition database.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics

Contributors: Iosifidis, A., Tefas, A., Pitas, I.

Number of pages: 13

Pages: 1445-1457
Publication date: Jun 2013
Peer-reviewed: Yes

Publication information

Journal: Signal Processing
Volume: 93
Issue number: 6
ISSN (Print): 0165-1684
Ratings:

Scopus rating (2013): CiteScore 5.1 SJR 0.909 SNIP 2.244

Original language: English

ASJC Scopus subject areas: Electrical and Electronic Engineering, Control and Systems Engineering, Software, Signal Processing, Computer Vision and Pattern Recognition

Keywords: Action recognition, Action volumes, Cluster discriminant analysis, Fuzzy vector quantization

DOIs:

10.1016/j.sigpro.2012.08.015

Source: Scopus

Source ID: 84875267100

Research output: Contribution to journal > Article > Scientific > peer-review

Dynamic action recognition based on dynemes and Extreme Learning Machine

In this paper, we propose a novel method that performs dynamic action classification by exploiting the effectiveness of the Extreme Learning Machine (ELM) algorithm for single hidden layer feedforward neural networks training. It involves data grouping and ELM based data projection in multiple levels. Given a test action instance, a neural network is trained by using labeled action instances forming the groups that reside to the test sample's neighborhood. The action instances involved in this procedure are, subsequently, mapped to a new feature space, determined by the trained network outputs. This procedure is performed multiple times, which are determined by the test action instance at hand, until only a single class is retained. Experimental results denote the effectiveness of the dynamic classification approach, compared to the static one, as well as the effectiveness of the ELM in the proposed dynamic classification setting.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics

Contributors: Iosifidis, A., Tefas, A., Pitas, I.

Number of pages: 9

Pages: 1890-1898

Publication date: 2013

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters

Volume: 34

Issue number: 15

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2013): CiteScore 4.8 SJR 0.768 SNIP 2.474

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Activity recognition, Dynamic classification, Extreme Learning Machine, Fuzzy vector quantization

DOIs:

10.1016/j.patrec.2012.10.019

Source: Scopus

Source ID: 84885069818

Research output: Contribution to journal > Article > Scientific > peer-review

Sparse extreme learning machine classifier exploiting intrinsic graphs

This paper presents an analysis of the recently proposed sparse extreme learning machine (S-ELM) classifier and describes an optimization scheme that can be used to calculate the network output weights. This optimization scheme exploits intrinsic graph structures in order to describe geometric data relationships in the so-called ELM space. Kernel formulations of the approach operating in ELM spaces of arbitrary dimensions are also provided. It is shown that the application of the optimization scheme exploiting geometric data relationships in the original ELM space is equivalent to the application of the original S-ELM to a transformed ELM space. The experimental results show that the incorporation of

geometric data relationships in S-ELM can lead to enhanced performance.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics, Aristotle University of Thessaloniki

Contributors: Iosifidis, A., Tefas, A., Pitas, I.

Number of pages: 5

Pages: 192-196

Publication date: 1 Nov 2015

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters

Volume: 65

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2015): CiteScore 5.1 SJR 0.95 SNIP 2.002

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Intrinsic graphs, Single-hidden layer neural networks, Sparse extreme learning machine

DOIs:

10.1016/j.patrec.2015.07.036

Source: Scopus

Source ID: 84940388000

Research output: Contribution to journal > Article > Scientific > peer-review

Nyström-based approximate kernel subspace learning

In this paper, we describe a method for the determination of a subspace of the feature space in kernel methods, which is suited to large-scale learning problems. Linear model learning in the obtained space corresponds to a nonlinear model learning process in the input space. Since the obtained feature space is determined only by exploiting properties of the training data, this approach can be used for generic nonlinear pattern recognition. That is, nonlinear data mapping can be considered to be a pre-processing step exploiting nonlinear relationships between the training data. Linear techniques can be subsequently applied in the new feature space and, thus, they can model nonlinear properties of the problem at hand. In order to appropriately address the inherent problem of kernel learning methods related to their time and memory complexities, we follow an approximate learning approach. We show that the method can lead to considerable operation speed gains and achieve very good performance. Experimental results verify our analysis.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research group: Video, Research Community on Data-to-Decision (D2D)

Contributors: Iosifidis, A., Gabbouj, M.

Number of pages: 8

Pages: 190-197

Publication date: Sep 2016

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition

ISSN (Print): 0031-3203

Ratings:

Scopus rating (2016): CiteScore 9 SJR 1.501 SNIP 3.005

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Kernel methods, Nonlinear pattern recognition, Nonlinear projection trick, Nyström approximation

DOIs:

10.1016/j.patcog.2016.03.018

Source: Scopus

Source ID: 85013223573

Research output: Contribution to journal > Article > Scientific > peer-review

Big Media Data Analysis

In this editorial a short introduction to the special issue on Big Media Data Analysis is given. The scope of this Editorial is to briefly present methodologies, tasks and applications of big media data analysis and to introduce the papers of the special issue. The special issue includes six papers that span various media analysis application areas like generic image description, medical image and video analysis, distance calculation acceleration and data collection.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Signal Processing, Research group: Multimedia Research Group - MRG, Aarhus Universitet, University of Milan Bicocca, Department of Informatics, Aristotle University of Thessaloniki

Contributors: Iosifidis, A., Tefas, A., Pitas, I., Gabbouj, M.

Number of pages: 4

Pages: 105-108

Publication date: 1 Nov 2017

Peer-reviewed: Yes

Publication information

Journal: Signal Processing: Image Communication

Volume: 59

ISSN (Print): 0923-5965

Ratings:

Scopus rating (2017): CiteScore 4.6 SJR 0.551 SNIP 1.512

Original language: English

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Electrical and Electronic Engineering

Keywords: Big Media Data, Data analytics, Deep learning, Machine learning, Statistical learning

DOIs:

10.1016/j.image.2017.10.004

Bibliographical note

EXT="Tefas, Anastasios"

Source: Scopus

Source ID: 85033445526

Research output: [Contribution to journal](#) › [Editorial](#) › [Scientific](#) › [peer-review](#)

The validity of using non-representative users in gaze communication research

Gaze-based interaction techniques have been investigated for the last two decades, and in many cases the evaluation of these has been based on trials with able-bodied users and conventional usability criteria, mainly speed and accuracy. The target user group of many of the gaze-based techniques investigated is, however, people with different types of physical disabilities. We present the outcomes of two studies that compare the performance of two groups of participants with a type of physical disability (one being cerebral palsy and the other muscular dystrophy) with that of a control group of able-bodied participants doing a task using a particular gaze interaction technique. One study used a task based on dwell-time selection, and the other used a task based on gaze gestures. In both studies, the groups of participants with physical disabilities performed significantly worse than the able-bodied control participants. We question the ecological validity of research into gaze interaction intended for people with physical disabilities that only uses able-bodied participants in evaluation studies without any testing using members of the target user population.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA), De Montfort University

Contributors: Istance, H., Vickers, S., Hyrskykari, A.

Number of pages: 4

Pages: 233-236

Publication date: 2012

Host publication information

Title of host publication: Proceedings - ETRA 2012: Eye Tracking Research and Applications Symposium

ISBN (Print): 9781450312257

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems

Keywords: assistive input devices, eye tracking, gaze communication, physically disabled user groups, representative users

DOIs:

10.1145/2168556.2168603

URLs:

<http://www.scopus.com/inward/record.url?scp=84862702657&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84862702657

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Full search equivalent fast block matching using orthonormal tree-structured haar transform

The goal of block matching is to find small parts (blocks) of an image that are similar to a given pattern (template). A lot of full search (FS) equivalent algorithms are based on transforms. However, the template size is limited to be a power-of-two. In this paper, we consider a fast block matching algorithm based on orthonormal tree-structured Haar transform (OTSHT) which makes it possible to use a template with arbitrary size. We evaluated the pruning performance, computational complexity, and design of tree. The pruning performance is compared to the algorithm based on orthonormal Haar transform (OHT).

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Tokyo Institute of Technology

Contributors: Ito, I., Egiazarian, K.

Number of pages: 6

Pages: 177-182

Publication date: 18 Oct 2017

Host publication information

Title of host publication: ISPA 2017 - 10th International Symposium on Image and Signal Processing and Analysis

Publisher: IEEE COMPUTER SOCIETY PRESS

ISBN (Electronic): 9781509040117

ASJC Scopus subject areas: Computational Theory and Mathematics, Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Signal Processing

DOIs:

10.1109/ISPA.2017.8073591

Bibliographical note

jufoid=57665

Source: Scopus

Source ID: 85037808698

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Opportunities and Challenges of Mobile Applications as "Tickets-to-Talk": A Scenario-Based User Study

This paper presents a scenario-based user study of mobile application concepts that would encourage interaction between people within close proximity. The scenarios demonstrate three themes of digital tickets-to-talk: informing who and what are around, augmenting self-expression, and online interaction encouraging physical interaction. Our interview study explored the opportunities and challenges of such applications in developing into further face-to-face interactions between strangers. Tickets that are related to activities that convey a solid intention that would lead to practical collaboration, such as playing sports or studying together, have the best potential to advance to meaningful face-to-face interaction. Augmenting selfexpression and online interaction encouraging physical interaction were found to have potential to create curiosity but seen less credible by our 42 interview participants to motivate face-to-face interaction between strangers. We conclude by discussing the potential of each theme of ticket-to-talk based on our findings as well as related literature.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Pervasive Computing, Augmented Human Activities (AHA)

Contributors: Jarusriboonchai, P., Olsson, T., Ojala, J., Väänänen-Vainio-Mattila, K.

Number of pages: 9

Pages: 89-97

Publication date: 2014

Host publication information

Title of host publication: Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia, MUM2014, November 25-28, 2014, Melbourne, Australia

Place of publication: New York, NY

Publisher: ACM

ISBN (Print): 978-1-4503-3304-7

Publication series

Name: International conference on mobile and ubiquitous multimedia

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Awareness system, Co-located interaction, Faceto-face interaction, Mobile technology, Scenarios, Storyboards , User experience, User study

DOIs:

10.1145/2677972.2677993

URLs:

<http://www.scopus.com/inward/record.url?scp=84943175665&partnerID=8YFLogxK> (Link to publication in Scopus)

Bibliographical note

Contribution: organisation=tie,FACT1=1
Portfolio EDEND: 2014-12-31
Publisher name: ACM

Source: researchoutputwizard

Source ID: 575

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Paraxial light beams in structured anisotropic media

We discuss the paraxial approximation for optical waves propagating in a uniaxial anisotropic medium inhomogeneously twisted on the plane normal to the wave vector, with the latter being parallel to one of the two principal axes normal to the optic axis. Such geometry implies a continuous power transfer between the ordinary and extraordinary components, regardless of the input beam polarization. We pinpoint that this peculiar feature, generalizable to any inhomogeneous linear birefringent material, strongly affects the application of the paraxial approximation due to the simultaneous presence of two different phase velocities. We eventually show that a local coordinate transformation permits a correct application of the paraxial approximation.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Photonics, Research group: Nonlinear Optics, Univ Porto, Universidade do Porto, Fac Med, Dept Med Imaging

Contributors: Jisha, C. P., Alberucci, A.

Number of pages: 6

Pages: 2019-2024

Publication date: 1 Nov 2017

Peer-reviewed: Yes

Publication information

Journal: Journal of the Optical Society of America A: Optics and Image Science, and Vision

Volume: 34

Issue number: 11

ISSN (Print): 1084-7529

Ratings:

Scopus rating (2017): CiteScore 3.5 SJR 0.687 SNIP 1.098

Original language: English

ASJC Scopus subject areas: Electronic, Optical and Magnetic Materials, Atomic and Molecular Physics, and Optics, Computer Vision and Pattern Recognition

DOIs:

10.1364/JOSAA.34.002019

Source: Scopus

Source ID: 85033368214

Research output: Contribution to journal › Article › Scientific › peer-review

Innovation challenges as a novel multidisciplinary learning platform

Innovation Challenges is a new course offered for the whole Tampere university community by Y-kampus entrepreneurship and innovation services, for the first time in fall 2019. Innovation Challenges offers practice-based cases that allow students to develop their creativity and problem-solving skills in a team. Learning is anchored in team coaching pedagogy, learning-by-doing attitude and entrepreneurial mindset. In this paper, we first describe the evolution that created a course called Innovation Challenges. Then, we describe course organization and the six challenges that student teams are currently solving.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Information and Knowledge Management, Research group: Business Data Research Group, Tampere University of Applied Sci., Tampere University
Contributors: Jokiniemi, S., Myllärniemi, J., Poranen, T., Vuorenmaa, M.
Number of pages: 4
Pages: 145-148
Publication date: 29 Jan 2020

Host publication information

Title of host publication: AcademicMindtrek 2020 - Proceedings of the 23rd International Academic Mindtrek Conference : January 2020, Tampere
Publisher: ACM
ISBN (Electronic): 9781450377744
ASJC Scopus subject areas: Software, Human-Computer Interaction, Computer Vision and Pattern Recognition, Computer Networks and Communications
Keywords: innovation, learning platform, multidisciplinary projects
DOIs:
10.1145/3377290.3377311

Bibliographical note

INT=comp,"Poranen, Timo"
Source: Scopus
Source ID: 85080863203
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Who is moving - User or device? Experienced quality of mobile 3D video in vehicles

'Viewing while commuting' is a typical use case for mobile video. However, experimental and behavioral influences of watching three-dimensional (3D) video in vibrating vehicles have not been widely researched. The goal of this study is 1) to explore the influence of video presentation modes (two-dimensional and stereoscopic 3D) on the quality of experience and 2) to understand the nature of the movement patterns that users perform to maintain an optimal viewing position while viewing videos on a mobile device in three commuting contexts and in a controlled laboratory environment. A hybrid method for quality evaluation was used for combining quantitative preference ratings, qualitative descriptions of quality, situational audio/video data-collection, and sensors. The high-quality and heterogeneous audiovisual stimuli were viewed on a mobile device equipped with a parallax barrier display. The results showed that the stereoscopic 3D (S3D) video presentation mode provided more satisfying quality of experience than the two-dimensional presentation mode in all studied contexts. To maintain an optimal viewing position in the vehicles, the users moved the device in their hands to the directions around the vertical and the horizontal axes in a leaned sitting position. This movement behavior was guided by the contexts but not by the quality, indicating the general importance of these results for mobile video viewing in vibrating vehicles.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Department of Pervasive Computing, Research area: User experience, Eindhoven University of Technology, Nokia
Contributors: Jumisko-Pyykkö, S., Markopoulos, P., Hannuksela, M. M.
Publication date: 16 Nov 2015

Host publication information

Title of host publication: ACE 2015 - 12th Advances in Computer Entertainment Technology Conference, Proceedings
Publisher: ACM
Article number: 13
ISBN (Electronic): 9781450338523
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: 3D, Experienced quality, Mobile video, Movement, Perception, Quasi-experiments
DOIs:
10.1145/2832932.2832948
Source: Scopus
Source ID: 84979759186
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Model for landmark highlighting in mobile web services

We introduce a model for landmark highlighting for pedestrian route guidance services for mobile devices. The model determines which landmarks are the most attractive based on their properties in the current context of user's orientation and the location on the route and highlights these landmarks on the mobile map. The attractiveness of a landmark is based on its visual, structural and semantic properties which are used for calculating the total attractiveness of a single landmark. This model was evaluated with voluntary users conducted in laboratory environment. Test subjects were shown images of street intersections from where they selected the most attractive and prominent landmarks in the route's context. We then compared these results with the landmarks selected by the model. The results show that landmarks highlighted by the model were the same ones that were selected by the participants as most salient landmarks.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA)

Contributors: Kallioniemi, P., Turunen, M.

Publication date: 2012

Host publication information

Title of host publication: Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, MUM 2012

Article number: 25

ISBN (Print): 9781450318150

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Human-Computer Interaction

Keywords: Landmarks, Mobile web services, Pedestrian route guidance

DOIs:

10.1145/2406367.2406398

URLs:

<http://www.scopus.com/inward/record.url?scp=84871605492&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84871605492

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Haptic feedback to gaze events

Eye tracking input often relies on visual and auditory feedback. Haptic feedback offers a previously unused alternative to these established methods. We describe a study to determine the natural time limits for haptic feedback to gazing events. The target is to determine how much time we can use to evaluate the user gazed object and decide if we are going to give the user a haptic notification on that object or not. The results indicate that it is best to get feedback faster than in 250 milliseconds from the start of fixation of an object. Longer delay leads to increase in incorrect associations between objects and the feedback. Delays longer than 500 milliseconds were confusing for the user.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA), School of Management (JKK)

Contributors: Kangas, J., Rantala, J., Majaranta, P., Isokoski, P., Raisamo, R.

Number of pages: 8

Pages: 11-18

Publication date: 2014

Host publication information

Title of host publication: Proceedings of the Symposium on Eye Tracking Research and Applications, ETRA 2014

Publisher: Association for Computing Machinery

ISBN (Print): 9781450327510

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems

Keywords: Gaze interaction, Gaze tracking, Haptic feedback

DOIs:

10.1145/2578153.2578154

URLs:

<http://www.scopus.com/inward/record.url?scp=84899691269&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84899691269

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

The Effect of Light Field Reconstruction and Angular Resolution Reduction on the Quality of Experience

The quality of visual contents displayed on 3D autostereoscopic displays—such as light field displays—essentially depend on factors that are not present in case of 3D stereoscopic or 2D ones, like angular resolution. A higher number of views in a given field of view enables a smoother, continuous motion parallax, but evidently requires more resources to transmit and display. However, in several cases a sufficiently high number of views might not even be available, thus light field reconstruction is required to increase the density of intermediate views. In this paper we introduce the results of a research aiming to measure the perceptual difference between light field reconstruction and different angular resolutions via a series of subjective image quality assessments. The analysis also calls attention to transmission requirements of content for light field displays.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Research group: 3D MEDIA, Robotic Vision Team, Kingston University, Holografika

Contributors: Kara, P. A., Kovacs, P. T., Vagharshakyan, S., Martini, M. G., Barsi, A., Balogh, T., Chuchvara, A., Chehaibi, A.

Number of pages: 6

Pages: 781-786

Publication date: 21 Apr 2017

Host publication information

Title of host publication: 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)

Publisher: IEEE

ISBN (Electronic): 9781509056989

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Radiology Nuclear Medicine and imaging, Computer Networks and Communications, Signal Processing

Keywords: Angular Resolution, Image Quality, Light Field Display, Light Field Reconstruction, Perceived Quality, Quality of Experience, View Synthesis

DOIs:

10.1109/SITIS.2016.128

Source: Scopus

Source ID: 85019236013

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Evaluation of visual object trackers on equirectangular panorama

Equirectangular (360° spherical) panorama is the most widely adopted format to store and broadcast virtual reality (VR) videos. Equirectangular projection provides a new challenge to adapt existing computer vision methods for the novel input type. In this work, we introduce a new dataset which consists of high quality equirectangular videos captured using a high-end VR camera (Nokia OZO). We also provide the original wide angle (8×195°) videos and densely annotated bounding boxes for evaluating object detectors and trackers. In this work, we introduce the dataset, compare state-of-the-art trackers for object tracking in equirectangular panorama and report detailed analysis of the failure cases which reveal potential factors to improve the existing visual object trackers for the new type of input.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Nokia Technologies

Contributors: Kart, U., Kämäräinen, J. K., Fan, L., Gabbouj, M.

Number of pages: 8

Pages: 25-32

Publication date: 2018

Host publication information

Title of host publication: VISIGRAPP 2018 - Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications

Volume: 5

Publisher: SCITEPRESS

ISBN (Electronic): 9789897582905

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Graphics and Computer-Aided Design, Artificial Intelligence

Keywords: 360°-video, Equirectangular, Tracking

DOIs:

10.5220/0006526200250032

Source: Scopus

Source ID: 85047804481

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Sparse approximations in complex domain based on BM3D modeling

In this paper the concept of sparsity for complex-valued variables is introduced in the following three types: directly in complex domain and for two real-valued pairs phase/amplitude and real/imaginary parts of complex variables. The nonlocal block-matching technique is used for sparsity implementation and filter design for each type of sparsity. These filters are complex domain generalizations of the Block Matching 3D collaborative (BM3D) filter based on the high-order singular value decomposition (HOSVD) in order to generate group-wise adaptive analysis/synthesis transforms. Complex domain denoising is developed and studied as a test-problem for comparison of the designed filters as well as the different types of sparsity modeling.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Signal Processing, Research group: Computational Imaging-CI

Contributors: Katkovnik, V., Ponomarenko, M., Egiazarian, K.

Number of pages: 13

Pages: 96-108

Publication date: 1 Dec 2017

Peer-reviewed: Yes

Publication information

Journal: Signal Processing

Volume: 141

ISSN (Print): 0165-1684

Ratings:

Scopus rating (2017): CiteScore 7.1 SJR 0.94 SNIP 1.974

Original language: English

ASJC Scopus subject areas: Control and Systems Engineering, Software, Signal Processing, Computer Vision and Pattern Recognition, Electrical and Electronic Engineering

Keywords: Block matching, Complex domain, Denoising, Elsevier article, Phase imaging, Sample document, Sparsity DOIs:

10.1016/j.sigpro.2017.05.032

Source: Scopus

Source ID: 85020311730

Research output: Contribution to journal > Article > Scientific > peer-review

Implications of audio and narration in the user experience design of virtual reality

Virtual reality (VR) is quickly gaining momentum as an immersive medium; however, there is much to learn about the design elements needed to create a positive experience. In this paper, we present the second wave of user testing of a journalistic and cultural VR experience that tells the story of a well-known artist through his art. The storytelling elements narration and ambient music were added to the initial prototype and tested in the field with 32 participants. Our results showed that the improvements produced a mostly positive user experience and shed light on what could be further improved in the case of our prototype, the field of immersive journalism, and VR used in the cultural context.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Pervasive Computing, Research area: User experience, Sanoma

Contributors: Kelling, C., Karhu, J., Kauhanen, O., Turunen, M., Väättäjä, H., Lindqvist, V.

Number of pages: 4

Pages: 258-261

Publication date: 10 Oct 2018

Host publication information

Title of host publication: Mindtrek 2018 - Proceedings of the 22nd International Academic Mindtrek Conference

Publisher: ACM

ISBN (Electronic): 9781450365895

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Cultural VR, Immersive Journalism, Museum, Storytelling

DOIs:

10.1145/3275116.3275153

Bibliographical note

INT=tie,"Kauhanen, Otto"

Source: Scopus

Source ID: 85056721502

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Community driven artificial intelligence development for robotics

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA)

Contributors: Kertész, C., Turunen, M.

Number of pages: 8

Pages: 3-10

Publication date: 2015

Host publication information

Title of host publication: Doctoral Consortium on Informatics in Control, Automation and Robotics, DCINCO 2015; in conjunction with the 12th International Conference on Informatics in Control, Automation and Robotics, ICINCO

Publisher: SCITEPRESS

ASJC Scopus subject areas: Control and Systems Engineering, Computer Vision and Pattern Recognition

URLs:

<http://www.scopus.com/inward/record.url?scp=84971280328&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84971280328

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Color-tone similarity of digital images

A color-tone similarity index (CSIM) between two color images is presented. CSIM is defined by a statistical analysis of cumulative histograms in a hue-oriented color space. It characterizes the color distributions, while the existing structural similarity index reflects the spatial structure involved with grayscale images. The behaviors of CSIM are checked by the comparisons of color code chips. Through an image quality assessment on TID2008, the correlation between CSIM and the mean opinion score was proved to be statistically significant.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: Vision, University of Niigata, KLab, Japan, Niigata University

Contributors: Kikuchi, H., Kataoka, S., Muramatsu, S., Huttunen, H.

Number of pages: 5

Pages: 393-397

Publication date: 2013

Host publication information

Title of host publication: 2013 IEEE International Conference on Image Processing, ICIP 2013 - Proceedings

ISBN (Print): 9781479923410

ASJC Scopus subject areas: Computer Vision and Pattern Recognition

Keywords: Image analysis, IQA, similarity, SSIM

DOIs:

10.1109/ICIP.2013.6738081

Source: Scopus

Source ID: 84897694822

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Implementation of a low-complexity low-latency arbitrary resampler on GPUs

Modern communication systems have data rates and sampling rates that are tightly coupled. Resampling is necessary in order to convert to some desired sampling rate, which is usually a multiple of the data rate. The resampling process is an integral part of transceiver systems and must be designed accurately and carefully. In this paper, we present a low complexity and low latency arbitrary resampling method based on graphics processing units (GPUs). Our proposed

flexible and all-software-based resampling method requires no precomputation of filters and yet yields high performance by taking advantage of unique features found in GPUs.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing Research Community (SPRC), University of Maryland, Department of Electrical and Computer Engineering

Contributors: Kim, S. C., Bhattacharyya, S. S.

Publication date: 21 Nov 2014

Host publication information

Title of host publication: 2014 IEEE Dallas Circuits and Systems Conference: Enabling an Internet of Things - From Sensors to Servers, DCAS 2014

Publisher: Institute of Electrical and Electronics Engineers Inc.

Article number: 6965333

ISBN (Electronic): 9781479959235

ASJC Scopus subject areas: Control and Systems Engineering, Electrical and Electronic Engineering, Computer Vision and Pattern Recognition

Keywords: Arbitrary resampling, DSP accelerator, GPU front-end receiver, GPU-based radio, sample rate conversion

DOIs:

10.1109/DCAS.2014.6965333

URLs:

<http://www.scopus.com/inward/record.url?scp=84918537027&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84918537027

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

An efficient GPU implementation of an arbitrary resampling polyphase channelizer

A channelizer is a part of a receiver front-end subsystem, commonly found in various communication systems, that separates different users or channels. A modern channelizer uses advantages of polyphase filter banks to process multiple channels at the same time, allowing down conversion, downsampling, and filtering all at the same time. However, due to limitations imposed by the structure and requirements of channelizers, their usage is limited and poses significant challenges due to inflexibility using conventional implementation techniques, which are intensively hardware-based. However, with advances in graphics processing unit (GPU) technology, we now have the potential to deliver high computational throughput along with the flexibility of software-based implementation. In this paper, we demonstrate how this potential can be exploited by presenting a novel GPU-based channelizer implementation. Our implementation incorporates methods for eliminating complex buffer managements and performing arbitrary resampling on all channels simultaneously. We also introduce the notion of simultaneously processing many channels as a high data rate parallel receiver system using blocks of threads in the GPU. The multi-channel, flexible, high-throughput, and arbitrary resampling characteristics of our GPU-based channelizer make it attractive for a variety of communication receiver applications.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing Research Community (SPRC), University of Maryland, Department of Electrical and Computer Engineering

Contributors: Kim, S. C., Plishker, W. L., Bhattacharyya, S. S.

Number of pages: 8

Pages: 231-238

Publication date: 2013

Host publication information

Title of host publication: DASIP 2013 - Proceedings of the 2013 Conference on Design and Architectures for Signal and Image Processing

Article number: 6661548

ISBN (Print): 9791092279016

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Hardware and Architecture, Signal Processing, Electrical and Electronic Engineering

Keywords: Arbitrary resampling, DSP accelerator, Front-end receiver, Polyphase channelizer, Sample rate conversion

URLs:

<http://www.scopus.com/inward/record.url?scp=84892642738&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84892642738

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

GPU-based acceleration of symbol timing recovery

This paper presents a novel implementation of graphics processing unit (GPU) based symbol timing recovery using polyphase interpolators to detect symbol timing error. Symbol timing recovery is a compute intensive procedure that detects and corrects the timing error in a coherent receiver. We provide optimal sample-time timing recovery using a maximum likelihood (ML) estimator to minimize the timing error. This is an iterative and adaptive system that relies on feedback, therefore, we present an accelerated implementation design by using a GPU for timing error detection (TED), enabling fast error detection by exploiting the 2D filter structure found in the polyphase interpolator. We present this hybrid/heterogeneous CPU and GPU architecture by computing a low complexity and low noise matched filter (MF) while simultaneously performing TED. We then compare the performance of the CPU vs. GPU based timing recovery for different interpolation rates to minimize the error and improve the detection by up to a factor of 35. We further improve the process by utilizing GPU optimization and performing block processing to improve the throughput even more, all while maintaining the lowest possible sampling rate.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing Research Community (SPRC), University of Maryland, Department of Electrical and Computer Engineering, Rice University

Contributors: Kim, S. C., Plishker, W. L., Bhattacharyya, S. S., Cavallaro, J. R.

Number of pages: 8

Pages: 273-280

Publication date: 2012

Host publication information

Title of host publication: DASIP 2012 - Proceedings of the 2012 Conference on Design and Architectures for Signal and Image Processing

Article number: 6385393

ISBN (Print): 9782953998726

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Hardware and Architecture, Signal Processing, Electrical and Electronic Engineering

Keywords: coherent receiver design, DSP accelerator, GPU, symbol timing recovery, synchronization

URLs:

<http://www.scopus.com/inward/record.url?scp=84872402791&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84872402791

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Demographic differences in accumulated types of capital in massively multiplayer online role-playing games

This paper examines how the demographic attributes and extra-game habits of players of a Massively Multiplayer Online Role-Playing Game (MMORPG) predict the accumulated capital of their avatars. An online survey (N=905) was conducted amidst the players of Final Fantasy XIV (FFXIV). Four types of capital were measured to map out the concrete and intangible resources of the avatars; social, economic, cultural and symbolic. The results show that weekly time spent playing the game is the strongest predictor of avatar capital and was associated with all types of capital. Time subscribed to the game was associated with cultural, economic, symbolic and bonding social capital. Social capital was found to be highest amongst both young and female players. Forum activity was associated with symbolic capital.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Communication Sciences, Research group: TUT Game Lab, Computing Sciences, Turun yliopisto

Contributors: Korkeila, H., Koivisto, J., Hamari, J.

Number of pages: 9

Pages: 74-82

Publication date: 29 Jan 2020

Host publication information

Title of host publication: AcademicMindtrek 2020 - Proceedings of the 23rd International Academic Mindtrek Conference : January 2020, Tampere

Publisher: ACM

ISBN (Electronic): 9781450377744

ASJC Scopus subject areas: Software, Human-Computer Interaction, Computer Vision and Pattern Recognition, Computer Networks and Communications

Keywords: avatar, capital, demographics, MMORPG

DOIs:

10.1145/3377290.3377302

Source: Scopus

Source ID: 85080910780

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

The visual object tracking VOT2013 challenge results

Visual tracking has attracted a significant attention in the last few decades. The recent surge in the number of publications on tracking-related problems have made it almost impossible to follow the developments in the field. One of the reasons is that there is a lack of commonly accepted annotated data-sets and standardized evaluation protocols that would allow objective comparison of different tracking methods. To address this issue, the Visual Object Tracking (VOT) workshop was organized in conjunction with ICCV2013. Researchers from academia as well as industry were invited to participate in the first VOT2013 challenge which aimed at single-object visual trackers that do not apply pre-learned models of object appearance (model-free). Presented here is the VOT2013 benchmark dataset for evaluation of single-object visual trackers as well as the results obtained by the trackers competing in the challenge. In contrast to related attempts in tracker benchmarking, the dataset is labeled per-frame by visual attributes that indicate occlusion, illumination change, motion change, size change and camera motion, offering a more systematic comparison of the trackers. Furthermore, we have designed an automated system for performing and evaluating the experiments. We present the evaluation protocol of the VOT2013 challenge and the results of a comparison of 27 trackers on the benchmark dataset. The dataset, the evaluation tools and the tracker rankings are publicly available from the challenge website (<http://votchallenge.net>).

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Research Community on Data-to-Decision (D2D), University of Ljubljana, Austrian Institute of Technology, University of Birmingham, Czech Technical University in Prague, Australian National University, DSTO, Sharif University of Technology, Nile University, Consorzio CREATE, University of South Australia, Vision and Sensing, ESTeM, University of Canberra, Panasonic RandD Center, CISP, University of Malaya, Eng. Design and Math., University of West England, Izmir Institute of Technology, Zhejiang University, Institute of Automation Chinese Academy of Sciences, Shanghai Institute of Ceramics Chinese Academy of Sciences, University of Surrey, Linköping University, Robotic Vision Team, Kingston University, NII, JFLI, NII

Contributors: Kristan, M., Pflugfelder, R., Leonardis, A., Matas, J., Porikli, F., Čehovin, L., Nebehay, G., Fernandez, G., Vojří, T., Gatt, A., Khajenezhad, A., Salahledin, A., Soltani-Farani, A., Zarezade, A., Petrosino, A., Milton, A., Bozorgtabar, B., Li, B., Chan, C. S., Heng, C., Ward, D., Kearney, D., Monekosso, D., Karaimer, H. C., Rabiee, H. R., Zhu, J., Gao, J., Xiao, J., Zhang, J., Xing, J., Huang, K., Lebeda, K., Cao, L., Maresca, M. E., Lim, M. K., ELHelw, M., Felsberg, M., Remagnino, P., Bowden, R., Goecke, R., Stolkin, R., Lim, S. Y. Y., Maher, S., Poullot, S., Wong, S., Satoh, S., Chen, W., Hu, W., Zhang, X., Li, Y., Niu, Z.

Number of pages: 14

Pages: 98-111

Publication date: 2013

Host publication information

Title of host publication: Proceedings - 2013 IEEE International Conference on Computer Vision Workshops, ICCVW 2013

Publisher: Institute of Electrical and Electronics Engineers Inc.

Article number: 6755885

ISBN (Print): 9781479930227

ASJC Scopus subject areas: Software, Computer Vision and Pattern Recognition

Keywords: Visual object tracking challenge, VOT2013

DOIs:

10.1109/ICCVW.2013.20

URLs:

<http://www.scopus.com/inward/record.url?scp=84897510119&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84897510119

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Goofy Mus, grumpy Mur and dirty Muf: Talking playful seats with personalities

The article discusses the concept of MurMur Moderators, talking playful seats designed to facilitate playful atmosphere and creativity at office environments. The concept of MurMur Moderators consists of five different personalities, grumpy Mur, goofy Mus, mellow Muh, sensitive Mut and shy Mum. The article describes the experiences and reactions to two personalities, Mus and Mur. Further, a sixth personality, Muf, consisting of rejected, provocative features is detailed. Consequently, the paper discusses play preferences, affordances and thresholds in connection to adult play. These will be the focus of future research by the authors.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Mathematical modelling with wide societal impact (MathImpact)
Contributors: Kultima, A., Nummenmaa, T., Tyni, H., Alha, K., Mayra, F.
Publication date: 11 Nov 2014

Host publication information

Title of host publication: ACE 2014 - 11th Advances in Computer Entertainment Technology Conference, Workshops Proceedings
Volume: 11-14-November-2014
Publisher: Association for Computing Machinery
Article number: a9
ISBN (Electronic): 9781450333146
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Adult play, Interactive furniture, Narrative, Personas, Playful office
DOIs:
10.1145/2693787.2693790
URLs:
<http://www.scopus.com/inward/record.url?scp=84962580216&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84962580216
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Local feature based unsupervised alignment of object class images

Alignment of objects is a predominant problem in part-based methods for visual object categorisation (VOC). These methods should learn the parts and their spatial variation, which is difficult for objects in arbitrary poses. A straightforward solution is to annotate images with a set of "object landmarks", but due to laborious manual annotation, semi-supervised methods requiring only a set of images and class labels are preferred. Recent state-of-the-art VOC methods utilise various approaches to align objects or otherwise compensate their geometric variation, but no explicit solution to the alignment problem with quantitative results can be found. The problem has been studied in the recent works related to "image congealing". The congealing methods, however, are based on image-based processing, and thus require moderate initial alignment and are sensitive to intra-class variation and background clutter. In this work, we define a local feature based algorithm to rigidly align object class images. Our algorithm is based on the standard VOC tools: local feature detectors and descriptors, correspondence based homography estimation, and random sample consensus (RANSAC) based spatial validation of local features. We first demonstrate how an intuitive feature matching approach works for simple classes, but fails for more complex ones. This is solved by a spatial scoring procedure which is the core element in the proposed method. Our method is compared to a state-of-the-art congealing method with realistic and difficult Caltech-101 and randomised Caltech-101 (r-Caltech-101) categories for which our method achieves clearly superior performance.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Research Community on Data-to-Decision (D2D), Lappeenranta University of Technology
Contributors: Lankinen, J., Kamarainen, J. K.
Publication date: 2011

Host publication information

Title of host publication: BMVC 2011 - Proceedings of the British Machine Vision Conference 2011
Publisher: British Machine Vision Association, BMVA
ASJC Scopus subject areas: Computer Vision and Pattern Recognition
DOIs:
10.5244/C25.107
URLs:
<http://www.scopus.com/inward/record.url?scp=84898425254&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84898425254
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Multi-sensor next-best-view planning as matroid-constrained submodular maximization

3D scene models are useful in robotics for tasks such as path planning, object manipulation, and structural inspection. We consider the problem of creating a 3D model using depth images captured by a team of multiple robots. Each robot selects a viewpoint and captures a depth image from it, and the images are fused to update the scene model. The process is repeated until a scene model of desired quality is obtained. Next-best-view planning uses the current scene model to select the next viewpoints. The objective is to select viewpoints so that the images captured using them improve the

quality of the scene model the most. In this letter, we address next-best-view planning for multiple depth cameras. We propose a utility function that scores sets of viewpoints and avoids overlap between multiple sensors. We show that multi-sensor next-best-view planning with this utility function is an instance of submodular maximization under a matroid constraint. This allows the planning problem to be solved by a polynomial-Time greedy algorithm that yields a solution within a constant factor from the optimal. We evaluate the performance of our planning algorithm in simulated experiments with up to 8 sensors, and in real-world experiments using two robot arms equipped with depth cameras.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Computing Sciences, Technical University Darmstadt, Max Planck Institute for Intelligent Systems, University of Hamburg

Contributors: Lauri, M., Pajarinen, J., Peters, J., Frintrop, S.

Number of pages: 8

Pages: 5323-5330

Publication date: 2020

Peer-reviewed: Yes

Publication information

Journal: IEEE Robotics and Automation Letters

Volume: 5

Issue number: 4

ISSN (Print): 2377-3766

Original language: English

ASJC Scopus subject areas: Control and Systems Engineering, Biomedical Engineering, Human-Computer Interaction, Mechanical Engineering, Computer Vision and Pattern Recognition, Computer Science Applications, Control and Optimization, Artificial Intelligence

Keywords: multi-robot systems, Reactive and sensor-based planning, RGB-D perception

DOIs:

10.1109/LRA.2020.3007445

Bibliographical note

EXT="Lauri, Mikko"

Source: Scopus

Source ID: 85090245712

Research output: Contribution to journal > Article > Scientific > peer-review

Long-term tracking through failure cases

Long term tracking of an object, given only a single instance in an initial frame, remains an open problem. We propose a visual tracking algorithm, robust to many of the difficulties which often occur in real-world scenes. Correspondences of edge-based features are used, to overcome the reliance on the texture of the tracked object and improve invariance to lighting. Furthermore we address long-term stability, enabling the tracker to recover from drift and to provide redetection following object disappearance or occlusion. The two-module principle is similar to the successful state-of-the-art long-term TLD tracker, however our approach extends to cases of low-textured objects. Besides reporting our results on the VOT Challenge dataset, we perform two additional experiments. Firstly, results on short-term sequences show the performance of tracking challenging objects which represent failure cases for competing state-of-the-art approaches. Secondly, long sequences are tracked, including one of almost 30000 frames which to our knowledge is the longest tracking sequence reported to date. This tests the re-detection and drift resistance properties of the tracker. All the results are comparable to the state-of-the-art on sequences with textured objects and superior on non-textured objects. The new annotated sequences are made publicly available.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Research Community on Data-to-Decision (D2D), University of Surrey, Czech Technical University in Prague

Contributors: Lebeda, K., Hadfield, S., Matas, J., Bowden, R.

Number of pages: 8

Pages: 153-160

Publication date: 2013

Host publication information

Title of host publication: Proceedings - 2013 IEEE International Conference on Computer Vision Workshops, ICCVW 2013

Publisher: Institute of Electrical and Electronics Engineers Inc.

Article number: 6755891

ISBN (Print): 9781479930227

ASJC Scopus subject areas: Software, Computer Vision and Pattern Recognition

Keywords: Computer vision, Edge, Line correspondence, Long-term tracking, Low texture, Visual tracking

DOIs:

10.1109/ICCVW.2013.26

URLs:

<http://www.scopus.com/inward/record.url?scp=84897541648&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84897541648

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Two patterns for minimizing human resources in a startup

In this paper, we describe two patterns that are part of a larger pattern language for software startup companies. These two particular patterns help startup companies to focus on the essential; the product itself and keeping their team intact and productive. In this way, the startup may operate with a sustainable team size.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Pervasive Computing, Research area: Software engineering

Contributors: Leppänen, M.

Publication date: 10 Apr 2014

Host publication information

Title of host publication: VikingPLoP 2014 Proceedings of the 8th Nordic Conference on Pattern Languages of Programs (VikingPLoP)

Publisher: ACM

Article number: 4

ISBN (Print): 9781450326605

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Lean startup, Organization, Organizational patterns, Patterns, Software engineering, Software product, Team

DOIs:

10.1145/2676680.2676686

Source: Scopus

Source ID: 84940021370

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Four patterns for internal startups

In this paper, we describe patterns that are meant for founding internal startups in a larger company. The patterns are part of a larger pattern language for software startup companies. The patterns presented here cover four main parts of an internal startup's life cycle starting from idea creation by enabling innovation with 20 Rule. The second pattern introduces an incubator phase, where the idea is validated to have a sensible problem and solution. This optimally leads to the creation of an internal startup, where resources are allocated to concretize the idea. With restricted resources such as a limited time, the internal startup may find a new Product-Market fit and offer a validated business opportunity for the parent company. This is concluded by the Exit decision by the parent company and ends the internal startup's life cycle.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Pervasive Computing, Research area: Software engineering

Contributors: Leppänen, M., Hokkanen, L.

Publication date: 8 Jul 2015

Host publication information

Title of host publication: Proceedings of the 20th European Conference on Pattern Languages of Programs, EuroPLoP 2015

Publisher: ACM

Article number: a5

ISBN (Electronic): 9781450338479

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Internal startup, Lean startup, Organization, Organizational patterns, Patterns

DOIs:

10.1145/2855321.2855327

Source: Scopus

Source ID: 84982784052

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Patterns for subsidiaries as innovation tools

In this paper, we describe two patterns for fostering innovative ideas in a company. The patterns originate from experiences in real companies. Innovations are crucial in opening up new business vistas for a company. Old business models for any company will wither as times change and, continuous innovation is needed. However, companies are geared for efficient execution of their current business, not for fostering new ideas. One way for innovation incubation is a subsidiary. A subsidiary typically has more freedom and risk-based incentives than an internal startup. To run a successful subsidiary, one must first decide when to Spin Off, then, how to run the Subsidiary and, finally, Merge and Scale the business, if feasible.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Pervasive Computing, Research area: User experience

Contributors: Leppänen, M., Hokkanen, L.

Publication date: 7 Apr 2016

Host publication information

Title of host publication: Proceedings of the 10th Travelling Conference on Pattern Languages of Programs, VikingPLoP 2016

Publisher: ACM

Article number: a7

ISBN (Electronic): 9781450342001

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Innovation, Internal startup, Lean, Startups

DOIs:

10.1145/3022636.3022643

Source: Scopus

Source ID: 85015703961

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Low power design methodology for signal processing systems using lightweight dataflow techniques

Dataflow modeling techniques facilitate many aspects of design exploration and optimization for signal processing systems, such as efficient scheduling, memory management, and task synchronization. The lightweight dataflow (LWDF) programming methodology provides an abstract programming model that supports dataflow-based design and implementation of signal processing hardware and software components and systems. Previous work on LWDF techniques has emphasized their application to DSP software implementation. In this paper, we present new extensions of the LWDF methodology for effective integration with hardware description languages (HDLs), and we apply these extensions to develop efficient methods for low power DSP hardware implementation. Through a case study of a deep neural network application for vehicle classification, we demonstrate our proposed LWDF-based hardware design methodology, and its effectiveness in low power implementation of complex signal processing systems.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Pervasive Computing, Signal Processing, Research group: Vision, Research area: Computer engineering, University of Maryland, Dept. of Electrical and Electronic Engineering, PolComIng - Information Engineering Unit, Department of Electrical and Computer Engineering

Contributors: Li, L., Fanni, T., Viitanen, T., Xie, R., Palumbo, F., Raffo, L., Huttunen, H., Takala, J., Bhattacharyya, S. S.

Number of pages: 8

Pages: 82-89

Publication date: 13 Feb 2017

Host publication information

Title of host publication: DASIP 2016 - Proceedings of the 2016 Conference on Design and Architectures for Signal and Image Processing

Publisher: IEEE COMPUTER SOCIETY PRESS

ISBN (Electronic): 9791092279153

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Hardware and Architecture, Signal Processing, Electrical and Electronic Engineering

DOIs:

10.1109/DASIP.2016.7853801

Bibliographical note

INT=tie,"Xie, Renjie"

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

A preliminary network analysis on steam game tags: Another way of understanding game genres

Video game genre classification has long been a focusing perspective in game studies domain. Despite the commonly acknowledged usefulness of genre classification, scholars in the game studies domain are yet to reach consensus on the game genre classification. On the other hand, Steam, a popular video game distribution platform, adopts the user-generated tag feature enabling players to describe and annotate video games based on their own understanding of genres. Despite the concern of the quality, the user-generated tags (game tags) provide an opportunity towards an alternative way of understanding video game genres based on the players' collective intelligence. Hence, in this study, we construct a network of game tags based on the co-occurrence of tags in games on Steam platform and analyze the structure of the network via centrality analysis and community detection. Such analysis shall provide an intuitive presentation on the distribution and connections of the game tags, which furthermore suggests a potential way of understanding the important tags that are commonly adopted and the main genres of video games.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Communication Sciences

Contributors: Li, X., Zhang, B.

Number of pages: 9

Pages: 65-73

Publication date: 29 Jan 2020

Host publication information

Title of host publication: AcademicMindtrek 2020 - Proceedings of the 23rd International Academic Mindtrek Conference : January 29-30, 2020, Tampere, Finland

Publisher: ACM

ISBN (Electronic): 9781450377744

ASJC Scopus subject areas: Software, Human-Computer Interaction, Computer Vision and Pattern Recognition, Computer Networks and Communications

Keywords: centrality, community detection, game tag, genre, modularity, network, steam, video game

DOIs:

10.1145/3377290.3377300

Bibliographical note

INT=coms,"Li, Xiaozhou"

Source: Scopus

Source ID: 85080924784

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

The use of advanced imaging technology in welfare technology solutions - Some ethical aspects

Advanced imaging technology with properties like a more realistic picture with extremely high resolution and new applications and branches like welfare technology where these properties are used also involves certain ethical challenges. The protection of vulnerable patients and the privacy of employees and third parties have not yet been discussed to any great extent but should be taken into account in designing, manufacturing and implementing the applications.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Satakunta University of Applied Sciences

Contributors: Lilja, K. K., Palomäki, J.

Number of pages: 4

Pages: 1-4

Publication date: 2 Feb 2018

Host publication information

Title of host publication: 3DTV-CON 2017 - 3D True Vision v2 : Research and Applications in Future 3D Media

Publisher: IEEE

ISBN (Electronic): 9781538616352

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Networks and Communications, Computer Vision and Pattern Recognition, Human-Computer Interaction, Electrical and Electronic Engineering

Keywords: 3D imaging, ethical, welfare technology

DOIs:

10.1109/3DTV.2017.8280396

Bibliographical note

jufoid=50006

Source: Scopus

Source ID: 85046368759

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Real-time human pose estimation with convolutional neural networks

In this paper, we present a method for real-time multi-person human pose estimation from video by utilizing convolutional neural networks. Our method is aimed for use case specific applications, where good accuracy is essential and variation of the background and poses is limited. This enables us to use a generic network architecture, which is both accurate and fast. We divide the problem into two phases: (1) pre-training and (2) finetuning. In pre-training, the network is learned with highly diverse input data from publicly available datasets, while in finetuning we train with application specific data, which we record with Kinect. Our method differs from most of the state-of-the-art methods in that we consider the whole system, including person detector, pose estimator and an automatic way to record application specific training material for finetuning. Our method is considerably faster than many of the state-of-the-art methods. Our method can be thought of as a replacement for Kinect in restricted environments. It can be used for tasks, such as gesture control, games, person tracking, action recognition and action tracking. We achieved accuracy of 96.8% (PCK@0.2) with application specific data.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Univ of Oulu, Aalto University

Contributors: Linna, M., Kannala, J., Rahtu, E.

Number of pages: 8

Pages: 335-342

Publication date: 2018

Host publication information

Title of host publication: VISIGRAPP 2018 - Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications

Volume: 5

Publisher: SCITEPRESS

ISBN (Electronic): 9789897582905

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Graphics and Computer-Aided Design, Artificial Intelligence

Keywords: Convolutional neural networks, Human pose estimation, Person detection

DOIs:

10.5220/0006624403350342

URLs:

<https://arxiv.org/pdf/1609.07420>

Source: Scopus

Source ID: 85047804818

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Game postmortems vs. developer Reddit AMAs: Computational analysis of developer communication

Postmortems and Reddit Ask Me Anything (AMA) threads represent communications of game developers through two different channels about their game development experiences, culture, processes, and practices. We carry out a quantitative text mining based comprehensive analysis of online available postmortems and AMA threads from game developers over multiple years. We find and analyze underlying topics from the postmortems and AMAs as well as their variation among the data sources and over time. The analysis is done based on structural topic modeling, a probabilistic modeling technique for text mining. The extracted topics reveal differing and common interests as well as their evolution of prevalence over time in the two text sources. We have found that postmortems put more emphasis on detail-oriented development aspects as well as technically-oriented game design problems whereas AMAs feature a wider variety of discussion topics that are related to a more general game development process, game-play and game-play experience related game design. The prevalences of the topics also evolve differently over time in postmortems versus AMAs.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Tampere University

Contributors: Lu, C., Peltonen, J., Nummenmaa, T.
Publication date: 26 Aug 2019

Host publication information

Title of host publication: Proceedings of the 14th International Conference on the Foundations of Digital Games, FDG 2019

Publisher: ACM

Editors: Khosmood, F., Pirker, J., Apperley, T., Deterding, S.

Article number: 22

ISBN (Electronic): 9781450372176

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Game development, Literature analysis, Postmortem analysis, Reddit, Text mining

DOIs:

10.1145/3337722.3337727

Bibliographical note

INT=comp,"Peltonen, Jaakko"

INT=comp,"Lu, Chien"

Source: Scopus

Source ID: 85072819939

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

'Aie-studio' - A pragmatist aesthetic approach for procedural sound design

This paper introduces the AIE-Studio (Audio Interfaces for Exploration), a modular dataflow patching library implemented with Pure Data. The AIE-Studio introduces new tools for procedural sound design through generative sonic and musical structures. Particular focus is on aesthetic experience. The designed modules allow versatile dataflow mapping through matrix routing system while also enabling the sound designer to influence generative processes of music creation. In particular, The AIE-Studio was used to create generative sonic and musical material in an embodied game-like application. In this paper we present key questions driving the research, theoretical background, research approach and the main development activities .

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA)

Contributors: Luhtala, M., Turunen, M., Hakulinen, J., Keskinen, T.

Publication date: 2013

Host publication information

Title of host publication: Proceedings of the 8th Audio Mostly: A Conference on Interaction with Sound, AM 2013 - In Cooperation with ACM SIGCHI

Publisher: Association for Computing Machinery

Article number: 7

ISBN (Print): 9781450326599

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Aesthetic experience, Artistic interfaces, Generative strategies, Procedural audio, Procedural sound design, Pure data, Sonic interaction design

DOIs:

10.1145/2544114.2544124

URLs:

<http://www.scopus.com/inward/record.url?scp=84898834763&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84898834763

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Systematic literature review on user logging in virtual reality

In this systematic literature review, we study the role of user logging in virtual reality research. By categorizing literature according to data collection methods and identifying reasons for data collection, we aim to find out how popular user logging is in virtual reality research. In addition, we identify publications with detailed descriptions about logging solutions. Our results suggest that virtual reality logging solutions are relatively seldom described in detail despite that many studies gather data by body tracking. Most of the papers gather data to witness something about a novel functionality or to compare different technologies without discussing logging details. The results can be used for scoping future virtual reality research.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Pervasive Computing, Research area: Software engineering
Contributors: Luoto, A.
Number of pages: 8
Pages: 110-117
Publication date: 10 Oct 2018

Host publication information

Title of host publication: Mindtrek 2018 - Proceedings of the 22nd International Academic Mindtrek Conference
Publisher: ACM
ISBN (Electronic): 9781450365895
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Systematic Literature Review, User Logging
Electronic versions:
Mindtrek_2018_Paper_PDF_file_14
DOIs:
10.1145/3275116.3275123
Source: Scopus
Source ID: 85056744675
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Log analysis of 360-degree video users via MQTT

Analyzing 360-degree video users is beneficial for 360-degree video application development. The analysis can be done with logged user data. In this paper, we argue that MQTT is a conventional technology for distributed logging of mobile 360-degree video users. MQTT not only saves resources also allows communication from the logging server to mobile clients in various networking conditions relatively easy. We constructed a proof of concept to show the feasibility of the approach. As log analysis examples, the proof of concept visualizes results of the most popular region of interest analysis and k-means clustering. The used research method is design science.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Computing Sciences
Contributors: Luoto, A.
Number of pages: 8
Pages: 130-137
Publication date: 2019

Host publication information

Title of host publication: ICGDA 2019 : Proceedings of the 2019 2nd International Conference on Geoinformatics and Data Analysis
Publisher: ACM
ISBN (Electronic): 978-1-4503-6245-0
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: 360-degree video, Component, Log Analysis, MQTT
Electronic versions:
ICSSE_2019_paper_5
DOIs:
10.1145/3318236.3318248
URLs:
<http://urn.fi/URN:NBN:fi:tuni-202009116968>
Source: Scopus
Source ID: 85066837109
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Stereoscopic video description for human action recognition

In this paper, a stereoscopic video description method is proposed that indirectly incorporates scene geometry information derived from stereo disparity, through the manipulation of video interest points. This approach is flexible and able to cooperate with any monocular low-level feature descriptor. The method is evaluated on the problem of recognizing complex human actions in natural settings, using a publicly available action recognition database of unconstrained stereoscopic 3D videos, coming from Hollywood movies. It is compared both against competing depth-aware approaches

and a state-of-the-art monocular algorithm. Experimental results denote that the proposed approach outperforms them and achieves state-of-the-art performance.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Research Community on Data-to-Decision (D2D), Aristotle University of Thessaloniki, Department of Informatics

Contributors: Mademlis, I., Iosifidis, A., Tefas, A., Nikolaidis, N., Pitas, I.

Publication date: 16 Jan 2015

Host publication information

Title of host publication: IEEE SSCI 2014 - 2014 IEEE Symposium Series on Computational Intelligence - CIMSIVP 2014: 2014 IEEE Symposium on Computational Intelligence for Multimedia, Signal and Vision Processing, Proceedings

Publisher: The Institute of Electrical and Electronics Engineers, Inc.

ISBN (Print): 9781479945047

ASJC Scopus subject areas: Artificial Intelligence, Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Human-Computer Interaction

DOIs:

10.1109/CIMSIVP.2014.7013263

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Information wall: Evaluation of a gesture-controlled public display

Public displays that allow users to interact with them through mid-air gestures are still relatively rare, as many applications rely on touch-based interaction. This paper introduces Information Wall, a gesture-controlled public information display that provides multi-user access to contextually relevant local information using remote pointing and mid-air gestures. The application has been studied in two settings: a lab-based user study and several short-term deployments. Based on our results, we present practical guidelines for gesture-controlled public display design.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA)

Contributors: Mäkelä, V., Heimonen, T., Luhtala, M., Turunen, M.

Number of pages: 4

Pages: 228-231

Publication date: 11 Nov 2014

Host publication information

Title of host publication: ACE 2014 - 11th Advances in Computer Entertainment Technology Conference, Proceedings

Volume: 2014-November

Publisher: Association for Computing Machinery

ISBN (Electronic): 9781450329453, 9781450331852, 9781450333047

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Gestures, Mid-air pointing, Pervasive displays, Public displays, User study

DOIs:

10.1145/2677972.2677998

URLs:

<http://www.scopus.com/inward/record.url?scp=84943142256&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84943142256

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Analysis of accommodation cues in holographic stereograms

The simplicity of the holographic stereogram (HS) makes it an attractive option in comparison to the more complex coherent computer generated hologram (CGH) methods. The cost of its simplicity is that the HS cannot accurately reconstruct deep scenes due to the lack of correct accommodation cues. The exact nature of the accommodation cues present in HSs, however, has not been investigated. In this paper, we provide analysis of the relation between the hologram sampling properties and the perceived accommodation response. The HS can be considered as a generator of a discrete light field (LF) and can thus be examined by considering the light ray oriented nature of the hologram diffracted light. We further support the analysis by employing a numerical reconstruction tool simulating the viewing process of the human eye. The simulation results demonstrate that HSs can provide accommodation cues depending on the choice of hologram segmentation size. It is further demonstrated that the accommodation response can be enhanced at the expense of loss in perceived spatial resolution.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Signal Processing, Research group: 3D MEDIA
Contributors: Mäkinen, J., Sahin, E., Gotchev, A.
Publication date: 1 Oct 2018

Host publication information

Title of host publication: 2018 - 3DTV-Conference : The True Vision - Capture, Transmission and Display of 3D Video, 3DTV-CON 2018
Publisher: IEEE
Article number: 8478586
ISBN (Electronic): 9781538661253
ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Networks and Communications, Computer Vision and Pattern Recognition, Human-Computer Interaction, Electrical and Electronic Engineering
Keywords: Accommodation, Holographic stereogram, Light field
Electronic versions:
3DTV-Con2018
DOIs:
10.1109/3DTV.2018.8478586
URLs:
<http://urn.fi/URN:NBN:fi:tuni-202003042531>

Bibliographical note

jufoid=50006
Source: Scopus
Source ID: 85056207484
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Identification of Parkinson's Disease Utilizing a Single Self-recorded 20-step Walking Test Acquired by Smartphone's Inertial Measurement Unit

Parkinson's disease (PD) is a degenerative and long-term disorder of the central nervous system, which often causes motor symptoms, e.g., tremor, rigidity, and slowness. Currently, the diagnosis of PD is based on patient history and clinical examination. Technology-derived decision support systems utilizing, for example, sensor-rich smartphones can facilitate more accurate PD diagnosis. These technologies could provide less obtrusive and more comfortable remote symptom monitoring. The recent studies showed that motor symptoms of PD can reliably be detected from data gathered via smartphones. The current study utilized an open-access dataset named 'mPower' to assess the feasibility of discriminating PD from non-PD by analyzing a single self-administered 20-step walking test. From this dataset, 1237 subjects (616 had PD) who were age and gender matched were selected and classified into PD and non-PD categories. Linear acceleration (ACC) and gyroscope (GYRO) were recorded by built-in sensors of smartphones. Walking bouts were extracted by thresholding signal magnitude area of the ACC signals. Features were computed from both ACC and GYRO signals and fed into a random forest classifier of size 128 trees. The classifier was evaluated deploying 100-fold cross-validation and provided an accumulated accuracy rate of 0.7 after 10k validations. The results show that PD and non-PD patients can be separated based on a single short-lasting self-administered walking test gathered by smartphones' built-in inertial measurement units.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Faculty of Biomedical Sciences and Engineering, Research group: Personal Health Informatics-PHI, Unit of Neurology, Satakunta Central Hospital
Contributors: Mehrang, S., Jauhiainen, M., Pietilä, J., Puustinen, J., Ruokolainen, J., Nieminen, H.
Number of pages: 4
Pages: 2913-2916
Publication date: 26 Oct 2018

Host publication information

Title of host publication: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2018
Volume: 2018-July
Publisher: Institute of Electrical and Electronics Engineers Inc.
Article number: 8512921
ISBN (Electronic): 9781538636466
ASJC Scopus subject areas: Signal Processing, Biomedical Engineering, Computer Vision and Pattern Recognition, Health Informatics

DOIs:

10.1109/EMBC.2018.8512921

Source: Scopus

Source ID: 85056600537

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Dominant Rotated Local Binary Patterns (DRLBP) for texture classification

In this paper, we present a novel rotation-invariant and computationally efficient texture descriptor called Dominant Rotated Local Binary Pattern (DRLBP). A rotation invariance is achieved by computing the descriptor with respect to a reference in a local neighborhood. A reference is fast to compute maintaining the computational simplicity of the Local Binary Patterns (LBP). The proposed approach not only retains the complete structural information extracted by LBP, but it also captures the complimentary information by utilizing the magnitude information, thereby achieving more discriminative power. For feature selection, we learn a dictionary of the most frequently occurring patterns from the training images, and discard redundant and non-informative features. To evaluate the performance we conduct experiments on three standard texture datasets: Outex12, Outex 10 and KTH-TIPS. The performance is compared with the state-of-the-art rotation invariant texture descriptors and results show that the proposed method is superior to other approaches.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research group: Computational Imaging-CI

Contributors: Mehta, R., Egiazarian, K.

Number of pages: 7

Pages: 16-22

Publication date: 2016

Peer-reviewed: Yes

Early online date: 30 Nov 2015

Publication information

Journal: Pattern Recognition Letters

Volume: 71

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2016): CiteScore 5.2 SJR 0.729 SNIP 1.678

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Feature Selection, KTH-TIPS, Local Binary Pattern (LBP), Outex, Rotation Invariance, Texture Classification

DOIs:

10.1016/j.patrec.2015.11.019

Source: Scopus

Source ID: 84951106920

Research output: Contribution to journal > Article > Scientific > peer-review

Image-Based Localization Using Hourglass Networks

In this paper, we propose an encoder-decoder convolutional neural network (CNN) architecture for estimating camera pose (orientation and location) from a single RGB-image. The architecture has a hourglass shape consisting of a chain of convolution and up-convolution layers followed by a regression part. The up-convolution layers are introduced to preserve the fine-grained information of the input image. Following the common practice, we train our model in end-to-end manner utilizing transfer learning from large scale classification data. The experiments demonstrate the performance of the approach on data exhibiting different lighting conditions, reflections, and motion blur. The results indicate a clear improvement over the previous state-of-the-art even when compared to methods that utilize sequence of test frames instead of a single frame.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Aalto University

Contributors: Melekhov, I., Ylioinas, J., Kannala, J., Rahtu, E.

Number of pages: 8

Pages: 870-877

Publication date: 19 Jan 2018

Host publication information

Title of host publication: 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017

Publisher: IEEE

ISBN (Electronic): 9781538610343

ASJC Scopus subject areas: Computer Science Applications, Computer Vision and Pattern Recognition

DOIs:

10.1109/ICCVW.2017.107

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

DGC-Net: Dense geometric correspondence network

This paper addresses the challenge of dense pixel correspondence estimation between two images. This problem is closely related to optical flow estimation task where Con-vNets (CNNs) have recently achieved significant progress. While optical flow methods produce very accurate results for the small pixel translation and limited appearance variation scenarios, they hardly deal with the strong geometric transformations that we consider in this work. In this paper, we propose a coarse-to-fine CNN-based framework that can leverage the advantages of optical flow approaches and extend them to the case of large transformations providing dense and subpixel accurate estimates. It is trained on synthetic transformations and demonstrates very good performance to unseen, realistic, data. Further, we apply our method to the problem of relative camera pose estimation and demonstrate that the model outperforms existing dense approaches.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Aalto University, Univ of Oulu, HCI e 486.1

Contributors: Melekhov, I., Tiulpin, A., Sattler, T., Pollefeys, M., Rahtu, E., Kannala, J.

Number of pages: 9

Pages: 1034-1042

Publication date: 4 Mar 2019

Host publication information

Title of host publication: 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019

Publisher: IEEE

ISBN (Electronic): 9781728119755

Publication series

Name: IEEE Winter Conference on Applications of Computer Vision

ISSN (Print): 1550-5790

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Science Applications

DOIs:

10.1109/WACV.2019.00115

Bibliographical note

jufoid=57596

Source: Scopus

Source ID: 85063572728

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Barriers for bridging interpersonal gaps: Three inspirational design patterns for increasing collocated social interaction

Positive face-to-face social encounters between strangers can strengthen the sense of community in modern urban environments. However, it is not always easy to initiate friendly encounters due to various inhibiting social norms. We present three inspirational design patterns for reducing inhibitions to interact with unfamiliar others. These abstractions are based on a broad design space review of concepts, encompassing examples across a range of scales, fields, media and forms. Each inspirational pattern is formulated as a response to a different challenge to initiating social interaction but all share an underlying similarity in offering varieties of barriers and filters that paradoxically also separate people. The patterns are "Closer Through Not Seeing"; "Closer Through Not Touching"; and "Minimize Encounter Duration". We believe these patterns can support designers, in understanding, articulating, and generating approaches to creating embodied interventions and systems that enable unacquainted people to interact.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Pervasive Computing, Research area: User experience, University of Southern Denmark

Contributors: Mitchell, R., Olsson, T.

Number of pages: 9

Pages: 2-10

Publication date: 26 Jun 2017

Host publication information

Title of host publication: C&T 2017 - 8th International Conference on Communities and Technologies, Conference Proceedings

Publisher: ACM

ISBN (Electronic): 9781450348546

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Collocated interaction, Face-to-face interaction, Social interaction design, pattern languages, embodied interaction

DOIs:

10.1145/3083671.3083697

Source: Scopus

Source ID: 85025125983

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Facilitating the first move: Exploring inspirational design patterns for aiding initiation of social encounters

Actualizing positive social encounters remains both a key ends and means in many activities to foster a sense of community. Initiating encounters between strangers typically requires facilitative activities or artefacts, such as icebreakers or tickets-to-talk. However, there is little understanding of which designs are effective and why, and the broad design space remains largely underexplored. We address this challenge by presenting five candidates for inspirational design patterns on signaling social intentions and identifying impediments that deter commencement of encounters. The principles result from an extensive review of design cases and public art installations. Through focus groups and expert interviews, we assessed the perceived applicability and social acceptance of the proposed patterns. Three new design principles relating to the risks of initiating an encounter emerged through analyzing participant responses. These articulations of possible approaches and pitfalls for increasing conviviality may broaden the repertoire of, and support discussion between designers and others concerned with collocated social interaction.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, University of Southern Denmark

Contributors: Mitchell, R., Olsson, T.

Number of pages: 12

Pages: 283-294

Publication date: 3 Jun 2019

Host publication information

Title of host publication: C&T 2019 - 9th International Conference on Communities and Technologies, Conference Proceedings

Publisher: ACM

Editors: Tellioglu, H., Cech, F.

ISBN (Electronic): 9781450371629

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Collocated interaction, Design patterns, Social encounters, Social encouragement, Social interaction design, Ticket-to-talk

DOIs:

10.1145/3328320.3328396

URLs:

<http://urn.fi/URN:NBN:fi:tuni-201908142894>

Source: Scopus

Source ID: 85067884637

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Controlled experiments comparing fault-tree-based safety analysis techniques

The capability to model dynamic aspects of safety-critical systems, such as sequence or stochastic dependence of events, is one important requirement for safety analysis techniques. State Event Fault Tree Analysis, Dynamic Fault Tree Analysis, and Fault Tree Analysis combined with Markov Chains Analysis have been developed to fulfill these requirements, but they are still not widely accepted and used in practice. In order to investigate the reasons behind this low usage, we conducted two controlled experiments. The goal of the experiments was to analyze and compare applicability and efficiency in State Event Fault Tree analysis versus Dynamic Fault Tree Analysis and Fault Tree Analysis combined with Markov Chains Analysis. The results of both experiments show that, notwithstanding the power of State Event Fault Tree Analysis, Dynamic Fault Tree Analysis is rated by participants as more applicable and is more efficient compared to State Event Fault Tree Analysis, which, in turn, is rated as more applicable but is less efficient than Fault Tree Analysis combined with Markov Chains Analysis. Two of the reasons investigated are the complexity of the notations used and the lack of tool support. Based on these results, we suggest strategies for enhancing State Event Fault Tree Analysis to

overcome its weaknesses and increase its applicability and efficiency in modeling dynamic aspects of safety-critical systems.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: University of Kaiserslautern, Former organisation of the author

Contributors: Mouaffo, A., Taibi, D., Jamboti, K.

Publication date: 2014

Host publication information

Title of host publication: 18th International Conference on Evaluation and Assessment in Software Engineering, EASE 2014

Publisher: Association for Computing Machinery (ACM)

Article number: a46

ISBN (Print): 9781450324762

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Controlled experiment, Dynamic fault tree, Fault tree analysis, Markov chain, Safety-analysis, Safety-critical systems, State event fault tree

DOIs:

10.1145/2601248.2601255

URLs:

<http://www.scopus.com/inward/record.url?scp=84905483353&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84905483353

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

On the Layer Selection in Small-Scale Deep Networks

Deep learning algorithms (in particular Convolutional Neural Networks, or CNNs) have shown their superiority in computer vision tasks and continue to push the state of the art in the most difficult problems of the field. However, deep models frequently lack interpretability. Current research efforts are often focused on increasingly complex and computationally expensive structures. These can be either handcrafted or generated by an algorithm, but in either case the specific choices of individual structural elements are hard to justify. This paper aims to analyze statistical properties of a large sample of small deep networks, where the choice of layer types is randomized. The limited representational power of such models forces them to specialize rather than generalize, resulting in several distinct structural patterns. Observing the empirical performance of structurally diverse weaker models thus allows for some practical insight into the connection between the data and the choice of suitable CNN architectures.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Research group: Multimedia Research Group - MRG

Contributors: Muravev, A., Raitoharju, J., Gabbouj, M.

Publication date: Nov 2018

Host publication information

Title of host publication: 2018 7th European Workshop on Visual Information Processing (EUVIP)

Publisher: IEEE

ISBN (Print): 978-1-5386-6898-6

ISBN (Electronic): 978-1-5386-6897-9

ASJC Scopus subject areas: Artificial Intelligence, Computer Vision and Pattern Recognition

Keywords: Multi-layer neural network, Supervised Learning, Pattern Analysis, Knowledge Representation

Electronic versions:

Layer Selection

DOIs:

10.1109/EUVIP.2018.8611738

URLs:

<http://urn.fi/URN:NBN:fi:tyy-201908262014>

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Ensembles of dense and dense sampling descriptors for the HEp-2 cells classification problem

The classification of Human Epithelial (HEp-2) cells images, acquired through Indirect Immunofluorescence (IIF) microscopy, is an effective method to identify staining patterns in patient sera. Indeed it can be used for diagnostic purposes, in order to reveal autoimmune diseases. However, the automated classification of IIF HEp-2 cell patterns

represents a challenging task, due to the large intra-class and the small inter-class variability. Consequently, recent HEP-2 cell classification contests have greatly spurred the development of new IIF image classification systems. Here we propose an approach for the automatic classification of IIF HEP-2 cell images by fusion of several texture descriptors by ensemble of support vector machines combined by sum rule. Its effectiveness is evaluated using the HEP-2 cells dataset used for the "Performance Evaluation of Indirect Immunofluorescence Image Analysis Systems" contest, hosted by the International Conference on Pattern Recognition in 2014: the accuracy on the testing set is 79.85%. The same dataset was used to test an ensemble of ternary-encoded local phase quantization descriptors, built by perturbation approaches: the accuracy on the training set is 84.16%. Finally, this ensemble was validated on 14 additional datasets, obtaining the best performance on 11 datasets. Our MATLAB code is available at <https://www.dei.unipd.it/node/2357>.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Electronics and Communications Engineering, Research group: Computational Biophysics and Imaging Group, BioMediTech, Integrated Technologies for Tissue Engineering Research (ITTE), Università degli Studi di Padova, Italy, University of Bologna

Contributors: Nanni, L., Lumini, A., dos Santos, F. L. C., Paci, M., Hyttinen, J.

Pages: 28-35

Publication date: 15 Oct 2016

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters

Volume: 82

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2016): CiteScore 5.2 SJR 0.729 SNIP 1.678

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Bag-of-features, Ensemble, HEP-2 cell classification, Machine learning, Support vector machine, Texture descriptors

Electronic versions:

Ensembles of dense and dense sampling descriptors for the HEP-2 cells classification problem. Embargo ended:

15/10/18

DOIs:

10.1016/j.patrec.2016.01.026

URLs:

<http://urn.fi/URN:NBN:fi:tty-201611294825>. Embargo ended: 15/10/18

Source: Scopus

Source ID: 84961195136

Research output: Contribution to journal > Article > Scientific > peer-review

Soft robotic gripper with compliant cell stacks for industrial part handling

Robot object grasping and handling requires accurate grasp pose estimation and gripper/end-effector design, tailored to individual objects. When object shape is unknown, cannot be estimated, or is highly complex, parallel grippers can provide insufficient grip. Compliant grippers can circumvent these issues through the use of soft or flexible materials that adapt to the shape of the object. This letter proposes a 3D printable soft gripper design for handling complex shapes. The compliant properties of the gripper enable contour conformation, yet offer tunable mechanical properties (i.e., directional stiffness). Objects that have complex shape, such as non-constant curvature, convex and/or concave shape can be grasped blind (i.e., without grasp pose estimation). The motivation behind the gripper design is handling of industrial parts, such as jet and Diesel engine components. (Dis)assembly, cleaning and inspection of such engines is a complex, manual task that can benefit from (semi-)automated robotic handling. The complex shape of each component, however, limits where and how it can be grasped. The proposed soft gripper design is tunable by compliant cell stacks that deform to the shape of the handled object. Individual compliant cells and cell stacks are characterized and a detailed experimental analysis of more than 600 grasps with seven different industrial parts evaluates the approach.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Automation Technology and Mechanical Engineering, Research group: Robotics and Automation

Contributors: Netzev, M., Angleraud, A., Pieters, R.

Number of pages: 8

Pages: 6821-6828

Publication date: 1 Oct 2020

Peer-reviewed: Yes

Publication information

Journal: IEEE Robotics and Automation Letters

Volume: 5

Issue number: 4

ISSN (Print): 2377-3766

Original language: English

ASJC Scopus subject areas: Control and Systems Engineering, Biomedical Engineering, Human-Computer Interaction, Mechanical Engineering, Computer Vision and Pattern Recognition, Computer Science Applications, Control and Optimization, Artificial Intelligence

Keywords: grasping, grippers and other end-effectors, Soft robotics

Electronic versions:

Soft Robotic Gripper With Compliant 2020

DOIs:

10.1109/LRA.2020.3020546

URLs:

<http://urn.fi/URN:NBN:fi:tuni-202009287146>

Source: Scopus

Source ID: 85091134388

Research output: [Contribution to journal](#) › [Article](#) › [Scientific](#) › [peer-review](#)

Social human-robot interaction in the wild: A workshop proposal for academic mindtrek 2020

This workshop will collect experts and stakeholders from all fields of human-robot interaction: both social and industrial applications and uses of robotics are of interest as far as they have human in the loop. The workshop will present recent and fully new research work in social HRI, including first results of a 3.5 month field trial and mixed-method study of a social robot Pepper in a shopping mall in Finland.

General information

Publication status: Published

MoE publication type: B3 Non-refereed article in conference proceedings

Organisations: Computing Sciences, Tampere University, VTT Technical Research Centre of Finland

Contributors: Niemelä, M., Ahtinen, A., Turunen, M.

Number of pages: 2

Pages: 168-169

Publication date: 29 Jan 2020

Host publication information

Title of host publication: AcademicMindtrek 2020 - Proceedings of the 23rd International Academic Mindtrek Conference : January 2020, Tampere

Publisher: ACM

ISBN (Electronic): 9781450377744

ASJC Scopus subject areas: Software, Human-Computer Interaction, Computer Vision and Pattern Recognition, Computer Networks and Communications

Keywords: collaborative robots, human-robot interaction, social robots

DOIs:

10.1145/3377290.3377320

Bibliographical note

INT=comp,"Turunen, Markku"

Source: Scopus

Source ID: 85080870105

Research output: [Chapter in Book/Report/Conference proceeding](#) › [Conference contribution](#) › [Scientific](#)

Deep Learning Case Study for Automatic Bird Identification

An automatic bird identification system is required for offshore wind farms in Finland. Indubitably, a radar is the obvious choice to detect flying birds, but external information is required for actual identification. We applied visual camera images as external data. The proposed system for automatic bird identification consists of a radar, a motorized video head and a single-lens reflex camera with a telephoto lens. A convolutional neural network trained with a deep learning algorithm is applied to the image classification. We also propose a data augmentation method in which images are rotated and converted in accordance with the desired color temperatures. The final identification is based on a fusion of parameters provided by the radar and the predictions of the image classifier. The sensitivity of this proposed system, on a dataset containing 9312 manually taken original images resulting in 2.44×10^6 augmented data set, is 0.9463 as an image classifier. The area under receiver operating characteristic curve for two key bird species is 0.9993 (the White-tailed Eagle) and 0.9496 (The Lesser Black-backed Gull), respectively. We proposed a novel system for automatic bird identification as a real world application. We demonstrated that our data augmentation method is suitable for image classification problem and it significantly increases the performance of the classifier.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Signal Processing, Research group: Data-analytics and Optimization, Mathematics
Contributors: Niemi, J., Tantt, J.
Number of pages: 15
Publication date: 29 Oct 2018
Peer-reviewed: Yes

Publication information

Journal: Applied Sciences (Switzerland)
Volume: 8
Issue number: 11
Article number: 2089
ISSN (Print): 2076-3417
Ratings:
Scopus rating (2018): SJR 0.379 SNIP 1.029
Original language: English
ASJC Scopus subject areas: Artificial Intelligence, Signal Processing, Computer Vision and Pattern Recognition
Keywords: Machine learning, Deep learning, Convolutional neural networks, Classification, data augmentation, intelligent surveillance systems
DOIs:
10.3390/app8112089
Research output: Contribution to journal > Article > Scientific > peer-review

Need to touch, wonder of discovery, and social capital: Experiences with interactive playful seats

In this article we present findings from a design experiment of MurMur Moderators, talking playful seats facilitating playful atmosphere and creativity at office environments. The article describes the design and technological composition of our two prototypes, and our experiences exposing the concept to audiences at science fairs and an office environment. This research has served as an exploratory design study, directing our focus to the seats as primary and secondary play objects with a distinct narrative. Our goal with the initial exposure was to first investigate preliminary audience reactions for the high level concept and how people interact with the prototype. This was then supplemented by testing the concept in an office environment. The data we have collected gives us insight on the seats as primary and secondary play objects and how users touch, discover and socialize.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Mathematical modelling with wide societal impact (MathImpact), RMIT University
Contributors: Nummenmaa, T., Tyni, H., Kultima, A., Alha, K., Holopainen, J.
Publication date: 16 Nov 2015

Host publication information

Title of host publication: ACE 2015 - 12th Advances in Computer Entertainment Technology Conference, Proceedings
Volume: 16-19-November-2015
Publisher: Association for Computing Machinery
Article number: 10
ISBN (Electronic): 9781450338523
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Arduino, Audio feedback, Design research, DIY, Game studies, Internet of things, Office play, Playful furniture, Raspberry Pi
DOIs:
10.1145/2832932.2832959
URLs:
<http://www.scopus.com/inward/record.url?scp=84979747766&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84979747766
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Social features in hybrid board game marketing material

This paper identifies 7 key social features which appear in the marketing and promotional material of hybrid board games. The features are identified by exploring sources such as game websites and game boxes of 13 hybrid board game products. The material is analyzed in order to determine how social features related to hybrid game features are presented. As a result of the analysis, it became apparent that there are certain key social features which are presented

as being important to players. The knowledge generated in this work acts as a view to how the industry sees hybridity in games as a tool for supporting social interaction, and how the industry wants to message it to consumers when they explore promotional material. The identified key social features can also be used as design knowledge for developing new games, as they give insight into popular social features in hybrid board games.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Computing Sciences, Tampere University
Contributors: Nummenmaa, T., Kankainen, V.
Publication date: 26 Aug 2019

Host publication information

Title of host publication: Proceedings of the 14th International Conference on the Foundations of Digital Games, FDG 2019
Publisher: ACM
Editors: Khosmood, F., Pirker, J., Apperley, T., Deterding, S.
Article number: 67
ISBN (Electronic): 9781450372176
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Board games, Design, Hybrid games, Marketing
DOIs:
10.1145/3337722.3341864

Bibliographical note

INT=comp,"Kankainen, Vill"
Source: Scopus
Source ID: 85072820010
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Designing for experiences with socially interactive robots

Socially interactive technologies are emerging as one of the predominant technologies of the future. In this workshop, we aim to discuss the emerging field of Social Robotic technologies with a particular focus on interaction design methodologies used in the design process. The workshop will investigate how researchers have approached designing social robots and what we can learn from the interaction design field for future designs. The main activities of the workshop will encompass two interactive sessions and a discussion panel on approaches to inspire the design of socially interactive robots. In particular, we focus on experience-driven design methods involving rituals and memorable experiences with social robots.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Pervasive Computing, Research area: User experience, Uppsala University, Strate School of Design
Contributors: Obaid, M., Kaipainen, K., Ocnarescu, I., Ahtinen, A.
Number of pages: 4
Pages: 948-951
Publication date: 29 Sep 2018

Host publication information

Title of host publication: NordiCHI 2018 : Revisiting the Life Cycle - Proceedings of the 10th Nordic Conference on Human-Computer Interaction
Publisher: ACM
ISBN (Electronic): 9781450364379
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Design, Social, Social Robot, Technology, User Experience
Electronic versions:
Obaid-et-al_NordiCHI2018_Designing-experiences_accepted-version (002)
DOIs:
10.1145/3240167.3240257
URLs:
<http://urn.fi/URN:NBN:fi:tuni-201910033689>
Source: Scopus
Source ID: 85056571102

Probabilistic approach to physical object disentangling

Physically disentangling entangled objects from each other is a problem encountered in waste segregation or in any task that requires disassembly of structures. Often there are no object models, and especially with cluttered irregularly shaped objects, the robot cannot create a model of the scene due to occlusion. One of our key insights is that based on previous sensory input we are only interested in moving an object out of the disentanglement around obstacles. That is, we only need to know where the robot can successfully move in order to plan the disentangling. Due to the uncertainty we integrate information about blocked movements into a probability map. The map defines the probability of the robot successfully moving to a specific configuration. Using as cost the failure probability of a sequence of movements we can then plan and execute disentangling iteratively. Since our approach circumvents only previously encountered obstacles, new movements will yield information about unknown obstacles that block movement until the robot has learned to circumvent all obstacles and disentangling succeeds. In the experiments, we use a special probabilistic version of the Rapidly exploring Random Tree (RRT) algorithm for planning and demonstrate successful disentanglement of objects both in 2-D and 3-D simulation, and, on a KUKA LBR 7-DOF robot. Moreover, our approach outperforms baseline methods.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Computing Sciences, Technical University Darmstadt, University of Lincoln, Max Planck Institute for Intelligent Systems

Contributors: Pajarinen, J., Arenz, O., Peters, J., Neumann, G.

Number of pages: 8

Pages: 5510-5517

Publication date: 2020

Peer-reviewed: Yes

Publication information

Journal: IEEE Robotics and Automation Letters

Volume: 5

Issue number: 4

ISSN (Print): 2377-3766

Original language: English

ASJC Scopus subject areas: Control and Systems Engineering, Biomedical Engineering, Human-Computer Interaction, Mechanical Engineering, Computer Vision and Pattern Recognition, Computer Science Applications, Control and Optimization, Artificial Intelligence

Keywords: Autonomous systems, collision avoidance, intelligent robots, path planning, probabilistic computing, waste recovery

DOIs:

10.1109/LRA.2020.3006789

Source: Scopus

Source ID: 85090290264

Research output: Contribution to journal › Article › Scientific › peer-review

Distant speech separation using predicted time-frequency masks from spatial features

Speech separation algorithms are faced with a difficult task of producing high degree of separation without containing unwanted artifacts. The time-frequency (T-F) masking technique applies a real-valued (or binary) mask on top of the signal's spectrum to filter out unwanted components. The practical difficulty lies in the mask estimation. Often, using efficient masks engineered for separation performance leads to presence of unwanted musical noise artifacts in the separated signal. This lowers the perceptual quality and intelligibility of the output. Microphone arrays have been long studied for processing of distant speech. This work uses a feed-forward neural network for mapping microphone array's spatial features into a T-F mask. Wiener filter is used as a desired mask for training the neural network using speech examples in simulated setting. The T-F masks predicted by the neural network are combined to obtain an enhanced separation mask that exploits the information regarding interference between all sources. The final mask is applied to the delay-and-sum beamformer (DSB) output. The algorithm's objective separation capability in conjunction with the separated speech intelligibility is tested with recorded speech from distant talkers in two rooms from two distances. The results show improvement in instrumental measure for intelligibility and frequency-weighted SNR over complex-valued non-negative matrix factorization (CNMF) source separation approach, spatial sound source separation, and conventional beamforming methods such as the DSB and minimum variance distortionless response (MVDR).

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research group: Audio research group

Contributors: Pertilä, P., Nikunen, J.

Number of pages: 10
Pages: 97-106
Publication date: 2015
Peer-reviewed: Yes

Publication information

Journal: Speech Communication
Volume: 68
ISSN (Print): 0167-6393
Ratings:

Scopus rating (2015): CiteScore 4.1 SJR 0.49 SNIP 1.612

Original language: English

ASJC Scopus subject areas: Modelling and Simulation, Computer Science Applications, Computer Vision and Pattern Recognition, Software, Communication, Linguistics and Language, Language and Linguistics

Keywords: Beamforming, Microphone arrays, Neural networks, Speech separation, Time-frequency masking

DOIs:

10.1016/j.specom.2015.01.006

Source: Scopus

Source ID: 84923277715

Research output: Contribution to journal > Article > Scientific > peer-review

Image database TID2013: Peculiarities, results and perspectives

This paper describes a recently created image database, TID2013, intended for evaluation of full-reference visual quality assessment metrics. With respect to TID2008, the new database contains a larger number (3000) of test images obtained from 25 reference images, 24 types of distortions for each reference image, and 5 levels for each type of distortion. Motivations for introducing 7 new types of distortions and one additional level of distortions are given; examples of distorted images are presented. Mean opinion scores (MOS) for the new database have been collected by performing 985 subjective experiments with volunteers (observers) from five countries (Finland, France, Italy, Ukraine, and USA). The availability of MOS allows the use of the designed database as a fundamental tool for assessing the effectiveness of visual quality. Furthermore, existing visual quality metrics have been tested with the proposed database and the collected results have been analyzed using rank order correlation coefficients between MOS and considered metrics. These correlation indices have been obtained both considering the full set of distorted images and specific image subsets, for highlighting advantages and drawbacks of existing, state of the art, quality metrics. Approaches to thorough performance analysis for a given metric are presented to detect practical situations or distortion types for which this metric is not adequate enough to human perception. The created image database and the collected MOS values are freely available for downloading and utilization for scientific purposes.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research group: Computational Imaging-CI, Research group: Algebraic and Algorithmic Methods in Signal Processing AAMSP, Signal Processing Research Community (SPRC), National Aerospace University, Dept of Transmitters, Receivers and Signal Processing, University of Rennes 1 - IETR, Media Communications Lab, USC Viterbi School of Engineering

Contributors: Ponomarenko, N., Jin, L., Ieremeiev, O., Lukin, V., Egiazarian, K., Astola, J., Vozel, B., Chehdi, K., Carli, M., Battisti, F., Jay Kuo, C. C.

Number of pages: 21

Pages: 57-77

Publication date: 1 Jan 2015

Peer-reviewed: Yes

Publication information

Journal: Signal Processing: Image Communication

Volume: 30

ISSN (Print): 0923-5965

Ratings:

Scopus rating (2015): CiteScore 4 SJR 0.532 SNIP 1.413

Original language: English

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Signal Processing, Software, Electrical and Electronic Engineering

Keywords: Image denoising, Image lossy compression, Image visual quality metrics

DOIs:

10.1016/j.image.2014.10.009

URLs:

<http://www.scopus.com/inward/record.url?scp=84919839405&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84919839405

Research output: Contribution to journal › Article › Scientific › peer-review

Revisiting gray pixel for statistical illumination estimation

We present a statistical color constancy method that relies on novel gray pixel detection and mean shift clustering. The method, called Mean Shifted Grey Pixel – MSGP, is based on the observation: true-gray pixels are aligned towards one single direction. Our solution is compact, easy to compute and requires no training. Experiments on two real-world benchmarks show that the proposed approach outperforms state-of-the-art methods in the camera-agnostic scenario. In the setting where the camera is known, MSGP outperforms all statistical methods.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Czech Technical University in Prague, Intel Finland

Contributors: Qian, Y., Pertuz, S., Nikkanen, J., Kämäräinen, J., Matas, J.

Number of pages: 11

Pages: 36-46

Publication date: 2019

Host publication information

Title of host publication: VISIGRAPP 2019 - Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications

Publisher: SCITEPRESS

Editors: Kerren, A., Hurter, C., Braz, J.

ISBN (Electronic): 9789897583544

ASJC Scopus subject areas: Computer Science Applications, Computer Vision and Pattern Recognition, Computer Graphics and Computer-Aided Design

Keywords: Color Constancy, Gray Pixel, Illumination Estimation

Electronic versions:

VISAPP_2019_201

DOIs:

10.5220/0007406900360046

URLs:

<http://urn.fi/URN:NBN:fi:ty-201908282042>

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Fast fourier color constancy and grayness index for ISPA illumination estimation challenge

We briefly introduce two submissions to the Illumination Estimation Challenge, in the Int'l Workshop on Color Vision, affiliated to the 11th Int'l Symposium on Image and Signal Processing and Analysis. The fourier-transform-based submission is ranked 3rd, and the statistical Gray-pixel-based one ranked 6th.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, South China University of Technology

Contributors: Qian, Y., Chen, K., Yu, H.

Number of pages: 3

Pages: 352-354

Publication date: 17 Oct 2019

Host publication information

Title of host publication: ISPA 2019 - 11th International Symposium on Image and Signal Processing and Analysis

Publisher: IEEE

Editors: Loncaric, S., Bregovic, R., Carli, M., Subasic, M.

ISBN (Electronic): 9781728131405

Publication series

Name: International Symposium on Image and Signal Processing and Analysis, ISPA

Volume: 2019-September

ISSN (Print): 1845-5921

ISSN (Electronic): 1849-2266

ASJC Scopus subject areas: Computational Theory and Mathematics, Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Color constancy, FFCC, Gray pixel, Illumination

DOIs:

10.1109/ISPA.2019.8868451

Bibliographical note

EXT="Chen, Ke"

Source: Scopus

Source ID: 85074428933

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Tut MUVIS image retrieval system proposal for MSR-Bing challenge 2014

This paper presents our system designed for MSR-Bing Image Retrieval Challenge @ ICME 2014. The core of our system is formed by a text processing module combined with a module performing PCA-assisted perceptron regression with random sub-space selection ($P^2R^2S^2$). $P^2R^2S^2$ uses Over-Feat features as a starting point and transforms them into more descriptive features via unsupervised training. The relevance score for each query-image pair is obtained by comparing the transformed features of the query image and the relevant training images. We also use a face bank, duplicate image detection, and optical character recognition to boost our evaluation accuracy. Our system achieves 0.5099 in terms of DCG_{25} on the development set and 0.5116 on the test set.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: Video, Tampere University of Technology, Research Community on Data-to-Decision (D2D)

Contributors: Raitoharju, J., Zhang, H., Ozan, E. C., Waris, M. A., Faisal, M., Cao, G., Roininen, M., Ahmad, I., Shetty, R., Uhlmann, S., Samiee, K., Kiranyaz, S., Gabbouj, M.

Number of pages: 6

Pages: 1-6

Publication date: 3 Sep 2014

Host publication information

Title of host publication: IEEE International Conference on Multimedia and Expo, ICME 2014, Chengdu, China, July 14-18, 2014

Place of publication: Piscataway

Publisher: Institute of Electrical and Electronics Engineers IEEE

ISBN (Print): 9781479947171

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Human-Computer Interaction

Keywords: Data Partitioning, Face Bank, Image Retrieval, Relevance Evaluation

DOIs:

10.1109/ICMEW.2014.6890600

URLs:

<http://www.scopus.com/inward/record.url?scp=84937152405&partnerID=8YFLogxK> (Link to publication in Scopus)

Bibliographical note

Contribution: organisation=sgn,FACT1=1
Portfolio EDEND: 2014-09-25

Source: researchoutputwizard

Source ID: 1331

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Benchmark database for fine-grained image classification of benthic macroinvertebrates

Managing the water quality of freshwaters is a crucial task worldwide. One of the most used methods to biomonitor water quality is to sample benthic macroinvertebrate communities, in particular to examine the presence and proportion of certain species. This paper presents a benchmark database for automatic visual classification methods to evaluate their ability for distinguishing visually similar categories of aquatic macroinvertebrate taxa. We make publicly available a new database, containing 64 types of freshwater macroinvertebrates, ranging in number of images per category from 7 to 577. The database is divided into three datasets, varying in number of categories (64, 29, and 9 categories). Furthermore, in order to accomplish a baseline evaluation performance, we present the classification results of Convolutional Neural Networks (CNNs) that are widely used for deep learning tasks in large databases. Besides CNNs, we experimented with several other well-known classification methods using deep features extracted from the data.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Signal Processing, Qatar University, University of Jyväskylä, Finnish Environment Institute
Contributors: Raitoharju, J., Riabchenko, E., Ahmad, I., Iosifidis, A., Gabbouj, M., Kiranyaz, S., Tirronen, V., Ärje, J., Kärkkäinen, S., Meissner, K.
Number of pages: 11
Pages: 73-83
Publication date: 1 Oct 2018
Peer-reviewed: Yes

Publication information

Journal: Image and Vision Computing

Volume: 78

ISSN (Print): 0262-8856

Ratings:

Scopus rating (2018): CiteScore 4.9 SJR 0.633 SNIP 1.655

Original language: English

ASJC Scopus subject areas: Signal Processing, Computer Vision and Pattern Recognition

Keywords: Benthic macroinvertebrates, Biomonitoring, Convolutional Neural Networks, Deep learning, Fine-grained classification

DOIs:

10.1016/j.imavis.2018.06.005

Source: Scopus

Source ID: 85052861257

Research output: Contribution to journal › Article › Scientific › peer-review

Gaussian mixture models for signal mapping and positioning

Maps of RSS from a wireless transmitter can be used for positioning or for planning wireless infrastructure. The RSS values measured at a single point are not always the same, but follow some distribution, which vary from point to point. In existing approaches in the literature this variation is neglected or its mapping requires making many measurements at every point, which makes the measurement collection very laborious. We propose to use GMs for modeling joint distributions of the position and the RSS value. The proposed model is more versatile than methods found in the literature as it models the joint distribution of RSS measurements and the location space. This allows us to model the distributions of RSS values in every point of space without making many measurement in every point. In addition, GMs allow us to compute conditional probabilities and posteriors of position in closed form. The proposed models can model any RSS attenuation pattern, which is useful for positioning in multifloor buildings. Our tests with WLAN signals show that positioning with the proposed algorithm provides accurate position estimates. We conclude that the proposed algorithm can provide useful information about distributions of RSS values for different applications.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Computing Sciences, Research group: Automation and Systems Theory, Aalto University, University of Liverpool, Universidad Antonio de Nebrija, Uppsala University

Contributors: Raitoharju, M., García-Fernández, F., Hostettler, R., Piché, R., Särkkä, S.

Publication date: 1 Mar 2020

Peer-reviewed: Yes

Early online date: 10 Oct 2019

Publication information

Journal: Signal Processing

Volume: 168

Article number: 107330

ISSN (Print): 0165-1684

Original language: English

ASJC Scopus subject areas: Control and Systems Engineering, Software, Signal Processing, Computer Vision and Pattern Recognition, Electrical and Electronic Engineering

Keywords: Gaussian mixtures, Indoor positioning, RSS, Signal mapping, Statistical modeling

DOIs:

10.1016/j.sigpro.2019.107330

URLs:

<http://urn.fi/URN:NBN:fi:tuni-201911085834>. Embargo ends: 10/10/21

Source: Scopus

Source ID: 85073693932

Research output: Contribution to journal › Article › Scientific › peer-review

Automatic word count estimation from daylong child-centered recordings in various language environments using language-independent syllabification of speech

Automatic word count estimation (WCE) from audio recordings can be used to quantify the amount of verbal communication in a recording environment. One key application of WCE is to measure language input heard by infants and toddlers in their natural environments, as captured by daylong recordings from microphones worn by the infants. Although WCE is nearly trivial for high-quality signals in high-resource languages, daylong recordings are substantially more challenging due to the unconstrained acoustic environments and the presence of near- and far-field speech. Moreover, many use cases of interest involve languages for which reliable ASR systems or even well-defined lexicons are not available. A good WCE system should also perform similarly for low- and high-resource languages in order to enable unbiased comparisons across different cultures and environments. Unfortunately, the current state-of-the-art solution, the LENA system, is based on proprietary software and has only been optimized for American English, limiting its applicability. In this paper, we build on existing work on WCE and present the steps we have taken towards a freely available system for WCE that can be adapted to different languages or dialects with a limited amount of orthographically transcribed speech data. Our system is based on language-independent syllabification of speech, followed by a language-dependent mapping from syllable counts (and a number of other acoustic features) to the corresponding word count estimates. We evaluate our system on samples from daylong infant recordings from six different corpora consisting of several languages and socioeconomic environments, all manually annotated with the same protocol to allow direct comparison. We compare a number of alternative techniques for the two key components in our system: speech activity detection and automatic syllabification of speech. As a result, we show that our system can reach relatively consistent WCE accuracy across multiple corpora and languages (with some limitations). In addition, the system outperforms LENA on three of the four corpora consisting of different varieties of English. We also demonstrate how an automatic neural network-based syllabifier, when trained on multiple languages, generalizes well to novel languages beyond the training data, outperforming two previously proposed unsupervised syllabifiers as a feature extractor for WCE.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Computing Sciences, Aalto University, Laboratoire de Sciences Cognitives et Psycholinguistique, Carnegie Mellon University, University of Manitoba, Max Planck Institute for Psycholinguistics, CONICET, Duke University
Contributors: Räsänen, O., Seshadri, S., Karadayi, J., Riebling, E., Bunce, J., Cristia, A., Metze, F., Casillas, M., Rosemberg, C., Bergelson, E., Soderstrom, M.

Number of pages: 18

Pages: 63-80

Publication date: 1 Oct 2019

Peer-reviewed: Yes

Publication information

Journal: Speech Communication

Volume: 113

ISSN (Print): 0167-6393

Ratings:

Scopus rating (2019): CiteScore 4.2 SJR 0.554 SNIP 1.297

Original language: English

ASJC Scopus subject areas: Software, Modelling and Simulation, Communication, Language and Linguistics, Linguistics and Language, Computer Vision and Pattern Recognition, Computer Science Applications

Keywords: Automatic syllabification, Daylong recordings, Language acquisition, Noise robustness, Word count estimation

Electronic versions:

1-s2.0-S0167639318304205-main

DOIs:

10.1016/j.specom.2019.08.005

URLs:

<http://urn.fi/URN:NBN:fi:tuni-201909173346>

Source: Scopus

Source ID: 85070952723

Research output: Contribution to journal > Article > Scientific > peer-review

User Positioning in mmW 5G Networks Using Beam-RSRP Measurements and Kalman Filtering

In this paper, we exploit the 3D-beamforming features of multiantenna equipment employed in fifth generation (5G) networks, operating in the millimeter wave (mmW) band, for accurate positioning and tracking of users. We consider sequential estimation of users' positions, and propose a two-stage extended Kalman filter (EKF) that is based on reference signal received power (RSRP) measurements. In particular, beamformed downlink (DL) reference signals (RSs) are transmitted by multiple base stations (BSs) and measured by user equipments (UEs) employing receive beamforming. The so-obtained beam-RSRP (BRSRP) measurements are reported to the BSs where the corresponding directions of departure (DoDs) are sequentially estimated by a novel EKF. Such angle estimates from multiple BSs are subsequently

fused on a central entity into 3D position estimates of UEs by means of another (second-stage) EKF. The proposed positioning scheme is scalable since the computational burden is shared among different network entities, namely transmission/reception points (TRPs) and 5G-NR Node B (gNB), and may be accomplished with the signalling currently specified for 5G. We assess the performance of the proposed algorithm on a realistic outdoor 5G deployment with a detailed ray tracing propagation model based on the METIS Madrid map. Numerical results with a system operating at 39 GHz show that sub-meter 3D positioning accuracy is achievable in future mmW 5G networks.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Electronics and Communications Engineering, Research group: Wireless Communications and Positioning, Huawei Technologies Oy (Finland) Co., Ltd

Contributors: Rastorgueva-Foi, E., Costa, M., Koivisto, M., Leppänen, K., Valkama, M.

Number of pages: 7

Pages: 1150-1156

Publication date: 5 Sep 2018

Host publication information

Title of host publication: 2018 21st International Conference on Information Fusion, FUSION 2018

Publisher: IEEE

Article number: 8455289

ISBN (Print): 978-1-5386-4330-3

ISBN (Electronic): 978-0-9964527-6-2

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Signal Processing, Statistics, Probability and Uncertainty, Instrumentation

Keywords: 5G networks, beamforming, direction-of-departure, extended Kalman filter, line-of-sight, localization, location-awareness, positioning, RSRP, tracking

Electronic versions:

User Positioning in mmW 5G Networks 2018

DOIs:

10.23919/ICIF.2018.8455289

URLs:

<http://urn.fi/URN:NBN:fi:tuni-202002262360>

Source: Scopus

Source ID: 85054063725

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Robotic process automation - Creating value by digitalizing work in the private healthcare?

Organizations are applying digitalization to the constantly increasing amounts of different organizational processes [2]. The healthcare sector is also changing and actively seeking better ways to enhance performance, especially in the private healthcare sector [7]. Automation of workflow processes, e.g., Robotic Process Automation (RPA), in organizations has been emerging as a solution to this demand [3, 4]. To meet this clear demand, automation of workflow processes in organizations has been a rising trend during the past few years [3]. We analyze the value creating functions of the RPA potential in the private healthcare industry sector, using modified Walter et al.'s function-oriented value analysis as our theoretical lens for identifying the potential of RPA.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Industrial and Information Management

Contributors: Ratia, M., Myllärniemi, J., Helander, N.

Number of pages: 6

Pages: 222-227

Publication date: 10 Oct 2018

Host publication information

Title of host publication: Mindtrek 2018 - Proceedings of the 22nd International Academic Mindtrek Conference

Publisher: ACM

ISBN (Electronic): 9781450365895

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Digitalization of knowledge work, Healthcare, Robotic Process Automation, Value creation

DOIs:

10.1145/3275116.3275129

Source: Scopus

Source ID: 85056714767

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Patterns for safety system bus architecture

Traditionally safety and controls systems have been strictly separated from each other. There are both benefits and liabilities in this approach. Thus, modern system employing control and safety system parts do not necessarily make a strict separation between these two elements of the system. Regardless of the degree of separation, the nodes belonging to either control or safety system may need to communicate with each other to implement the desired functionality. An increasing number of systems nowadays utilize a fieldbus to connect the distributed nodes of the system together. A time comes in the design process, when one needs to select the architecture of the physical fieldbus. That is, how and which nodes are connected? In this paper, two patterns to organize the fieldbus architecture are illustrated. In short, one which can separate the fieldbus between the safety and control system nodes or use a shared fieldbus between the nodes.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Research area: Information Systems in Automation, Automation and Hydraulic Engineering

Contributors: Rauhamäki, J.

Publication date: 7 Apr 2016

Host publication information

Title of host publication: Proceedings of the 10th Travelling Conference on Pattern Languages of Programs, VikingPLOP 2016

Publisher: ACM

Article number: a4

ISBN (Electronic): 9781450342001

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Communication, Control system, Fieldbus, Safety system, Separated, Shared

DOIs:

10.1145/3022636.3022640

Source: Scopus

Source ID: 85015687535

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Generative part-based Gabor object detector

Discriminative part-based models have become the approach for visual object detection. The models learn from a large number of positive and negative examples with annotated class labels and location (bounding box). In contrast, we propose a part-based generative model that learns from a small number of positive examples. This is achieved by utilizing "privileged information", sparse class-specific landmarks with semantic meaning. Our method uses bio-inspired complex-valued Gabor features to describe local parts. Gabor features are transformed to part probabilities by unsupervised Gaussian Mixture Model (GMM). GMM estimation is robustified for a small amount of data by a randomization procedure inspired by random forests. The GMM framework is also used to construct a probabilistic spatial model of part configurations. Our detector is invariant to translation, rotation and scaling. On part level invariance is achieved by pose quantization which is more efficient than previously proposed feature transformations. In the spatial model, invariance is achieved by mapping parts to an "aligned object space". Using a small number of positive examples our generative method performs comparably to the state-of-the-art discriminative method.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research group: Vision, Field robotics for efficient work sites (FIRE), Prostate cancer research center (PCRC), Lappeenranta University of Technology

Contributors: Riabchenko, E., Kämäräinen, J.

Number of pages: 8

Pages: 1-8

Publication date: 15 Dec 2015

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters

Volume: 68

Issue number: P1

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2015): CiteScore 5.1 SJR 0.95 SNIP 2.002

Original language: English

ASJC Scopus subject areas: Software, Artificial Intelligence, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Gabor feature, Gaussian mixture model, Generative learning, Object detection, Visual classification

DOIs:

10.1016/j.patrec.2015.08.004

Bibliographical note

EXT="Riabchenko, Ekaterina"

Source: Scopus

Source ID: 84941570575

Research output: Contribution to journal > Article > Scientific > peer-review

A comparison of security assurance support of agile software development methods

Agile methods increase the speed and reduce the cost of software projects; however, they have been criticized for lack of documentation, traditional quality control, and, most importantly, lack of security assurance - mostly due to their informal and self-organizing approach to software development. This paper clarifies the requirements for security assurance by using an evaluation framework to analyze the compatibility of established agile security development methods: XP, Scrum and Kanban, combined with Microsoft SDL security framework, against Finland's established national security regulation (Vahti). We also analyze the selected methods based on their role definitions, and provide some avenues for future research.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Managing digital industrial transformation (mDIT), University of Turku, Department of Information Technology

Contributors: Rindell, K., Hyrynsalmi, S., Leppänen, V.

Number of pages: 8

Pages: 61-68

Publication date: 25 Jun 2015

Host publication information

Title of host publication: Computer Systems and Technologies - 16th International Conference, CompSysTech 2015: Proceedings

Volume: 1008

Publisher: Association for Computing Machinery

ISBN (Electronic): 9781450333573

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: DESMET, Kanban, Scrum, SDL, Secure agile development, Security assurance, Vahti, XP

DOIs:

10.1145/2812428.2812431

Source: Scopus

Source ID: 84957689583

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Is Texture Denoising Efficiency Predictable?

Images of different origin contain textures, and textural features in such regions are frequently employed in pattern recognition, image classification, information extraction, etc. Noise often present in analyzed images might prevent a proper solution of basic tasks in the aforementioned applications and is worth suppressing. This is not an easy task since even the most advanced denoising methods destroy texture in a more or less degree while removing noise. Thus, it is desirable to predict the filtering behavior before any denoising is applied. This paper studies the efficiency of texture image denoising for different noise intensities and several filter types under different visual quality criteria (quality metrics). It is demonstrated that the most efficient existing filters provide very similar results. From the obtained results, it is possible to generalize and employ the prediction strategy earlier proposed for denoising techniques based on the discrete cosine transform. Accuracy of such a prediction is studied and the ways to improve it are considered. Some practical recommendations concerning a decision to undertake whether it is worth applying a filter are given.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Signal Processing, Research group: Computational Imaging-CI

Contributors: Rubel, O., Lukin, V., Abramov, S., Vozel, B., Pogrebnyak, O., Egiazarian, K.
Publication date: 2018
Peer-reviewed: Yes

Publication information

Journal: International Journal of Pattern Recognition and Artificial Intelligence

Volume: 32

Issue number: 1

Article number: 1860005

ISSN (Print): 0218-0014

Ratings:

Scopus rating (2018): CiteScore 2.2 SJR 0.304 SNIP 0.718

Original language: English

ASJC Scopus subject areas: Software, Computer Vision and Pattern Recognition, Artificial Intelligence

Keywords: image processing, noise suppression, Texture denoising, visual quality

Electronic versions:

S0218001418600054

DOIs:

10.1142/S0218001418600054

URLs:

<http://urn.fi/URN:NBN:fi:tty-201708241829>

Source: Scopus

Source ID: 85025804856

Research output: Contribution to journal > Article > Scientific > peer-review

Exploring the use of deprecated PHP releases in the wild internet: Still a LAMP issue?

Many web sites utilize deprecated software products that are no longer maintained by the associated software producers. This paper explores the question of whether an existing big data collection can be used to predict the likelihood of deprecated PHP releases based on different abstract components in modern web deployment stacks. Building on web intelligence, software security, and data-based industry rationales, the question is examined by focusing on the most popular domains in the contemporary web-facing Internet. Logistic regression is used for classification. Although statistical classification performance is modest, the results indicate that deprecated PHP releases are associated with Linux and other open source software components. Geographical variation is small. Besides these results, the paper contributes to the web intelligence research by evaluating the feasibility of existing big data collections for mass-scale fingerprinting.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: University of Turku, Department of Information Technology

Contributors: Ruohonen, J., Hyrynsalmi, S., Leppänen, V.

Publication date: 13 Jun 2016

Host publication information

Title of host publication: 6th International Conference on Web Intelligence, Mining and Semantics, WIMS 2016

Publisher: Association for Computing Machinery

Article number: 26

ISBN (Electronic): 9781450340564

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Cyber security, Patching, Release engineering, Web crawling

DOIs:

10.1145/2912845.2912851

Source: Scopus

Source ID: 84978522051

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Input magnitude data setting in error-reduction algorithm for one-dimensional discrete phase retrieval problem

In this paper we discuss how does the input magnitude data setting influence the behavior of error-reduction algorithm in the case of the one-dimensional discrete phase retrieval problem. We present experimental results related to the convergence or stagnation of the algorithm. We also discuss the issue of the zeros distribution of the solution, when the solution of the problem exists.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication
Organisations: Computing Sciences, FETTI, Technical University of Cluj-Napoca
Contributors: Rusu, C., Astola, J.
Publication date: 1 Jul 2019

Host publication information

Title of host publication: ISSCS 2019 - International Symposium on Signals, Circuits and Systems
Publisher: IEEE
Article number: 8801743
ISBN (Electronic): 9781728138961
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Hardware and Architecture, Signal Processing, Electrical and Electronic Engineering
DOIs:
10.1109/ISSCS.2019.8801743

Bibliographical note

EXT="Rusu, Corneliu"
Source: Scopus
Source ID: 85071848180
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Calculation of the scalar diffraction field from curved surfaces by decomposing the three-dimensional field into a sum of Gaussian beams

We present a local Gaussian beam decomposition method for calculating the scalar diffraction field due to a twodimensional field specified on a curved surface. We write the three-dimensional field as a sum of Gaussian beams that propagate toward different directions and whose waist positions are taken at discrete points on the curved surface. The discrete positions of the beam waists are obtained by sampling the curved surface such that transversal components of the positions form a regular grid. The modulated Gaussian window functions corresponding to Gaussian beams are placed on the transversal planes that pass through the discrete beam-waist position. The coefficients of the Gaussian beams are found by solving the linear system of equations where the columns of the system matrix represent the field patterns that the Gaussian beams produce on the given curved surface. As a result of using local beams in the expansion, we end up with sparse system matrices. The sparsity of the system matrices provides important advantages in terms of computational complexity and memory allocation while solving the system of linear equations.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Signal Processing Research Community (SPRC), Bilkent University
Contributors: Şahin, E., Onural, L.
Number of pages: 10
Pages: 527-536
Publication date: 2013
Peer-reviewed: Yes

Publication information

Journal: Journal of the Optical Society of America A: Optics Image Science and Vision
Volume: 30
Issue number: 3
ISSN (Print): 1084-7529
Ratings:
Scopus rating (2013): CiteScore 3.3 SJR 1.046 SNIP 1.331
Original language: English
ASJC Scopus subject areas: Atomic and Molecular Physics, and Optics, Electronic, Optical and Magnetic Materials, Computer Vision and Pattern Recognition
URLs:
<http://www.scopus.com/inward/record.url?scp=84875512966&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84875512966
Research output: Contribution to journal > Article > Scientific > peer-review

Scalar diffraction field calculation from curved surfaces via Gaussian beam decomposition

We introduce a local signal decomposition method for the analysis of three-dimensional (3D) diffraction fields involving curved surfaces. We decompose a given field on a two-dimensional curved surface into a sum of properly shifted and modulated Gaussian-shaped elementary signals. Then we write the 3D diffraction field as a sum of Gaussian beams, each of which corresponds to a modulated Gaussian window function on the curved surface. The Gaussian beams are

propagated according to a derived approximate expression that is based on the Rayleigh-Sommerfeld diffraction model. We assume that the given curved surface is smooth enough that the Gaussian window functions on it can be treated as written on planar patches. For the surfaces that satisfy this assumption, the simulation results show that the proposed method produces quite accurate 3D field solutions.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Signal Processing Research Community (SPRC), Bilkent University
Contributors: Şahin, E., Onural, L.
Number of pages: 11
Pages: 1459-1469
Publication date: 1 Jul 2012
Peer-reviewed: Yes

Publication information

Journal: Journal of the Optical Society of America A: Optics Image Science and Vision
Volume: 29
Issue number: 7
ISSN (Print): 1084-7529
Ratings:
Scopus rating (2012): CiteScore 3.2 SJR 1.065 SNIP 1.198
Original language: English
ASJC Scopus subject areas: Atomic and Molecular Physics, and Optics, Electronic, Optical and Magnetic Materials, Computer Vision and Pattern Recognition
DOIs:
10.1364/JOSAA.29.001459
URLs:
<http://www.scopus.com/inward/record.url?scp=84863743776&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84863743776
Research output: Contribution to journal › Article › Scientific › peer-review

Olfactory display prototype for presenting and sensing authentic and synthetic odors

The aim was to study if odors evaporated by an olfactory display prototype can be used to affect participants' cognitive and emotion-related responses to audio-visual stimuli, and whether the display can benefit from objective measurement of the odors. The results showed that odors and videos had significant effects on participants' responses. For instance, odors increased pleasantness ratings especially when the odor was authentic and the video was congruent with odors. The objective measurement of the odors was shown to be useful. The measurement data was classified with 100 % accuracy removing the need to speculate whether the odor presentation apparatus is working properly.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Pervasive Computing, Faculty of Biomedical Sciences and Engineering, Research group: Sensor Technology and Biomeasurements (STB), Research group: Micro and Nanosystems Research Group
Contributors: Salminen, K., Rantala, J., Isokoski, P., Lehtonen, M., Müller, P., Karjalainen, M., Väliäho, J., Kontunen, A., Nieminen, V., Leivo, J., Telembeci, A. A., Lekkala, J., Kallio, P., Surakka, V.
Number of pages: 5
Pages: 73-77
Publication date: 2 Oct 2018

Host publication information

Title of host publication: ICMI 2018 - Proceedings of the 2018 International Conference on Multimodal Interaction
Publisher: ACM
ISBN (Electronic): 9781450356923
ASJC Scopus subject areas: Computer Science Applications, Computer Vision and Pattern Recognition, Hardware and Architecture, Human-Computer Interaction
Keywords: Emotions, Multimodal interaction, Olfaction
Electronic versions:
olfactory-display-prototype
DOIs:
10.1145/3242969.3242999
URLs:

<http://urn.fi/URN:NBN:fi:tuni-201909233451>

Bibliographical note

INT=tut-bmt,"Nieminen, Ville"

Source: Scopus

Source ID: 85056660798

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Mixed reality with multimodal head-mounted pico projector

Many kinds of displays can be used for augmented reality (AR). Multimodal head-mounted pico projector is a concept, which is little explored for AR. It opens new possibilities for wearable displays. In this paper we present our proof-of-concept prototype of a multimodal head-mounted pico projector. Our main contributions are the display concept and some usage examples for it.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA), University of Tampere

Contributors: Sand, A., Rakkolainen, I.

Publication date: 2013

Host publication information

Title of host publication: Proceedings of the Virtual Reality International Conference on Laval Virtual, VRIC 2013

Article number: 14

ISBN (Print): 9781450318754

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Augmented reality, Mixed reality, Multimodality, Pico projector, Wearable displays

DOIs:

10.1145/2466816.2466831

URLs:

<http://www.scopus.com/inward/record.url?scp=84882277921&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84882277921

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Efficient Solving of Markov Decision Processes on GPUs Using Parallelized Sparse Matrices

Markov Decision Processes (MDPs) provide important capabilities for facilitating the dynamic adaptation of hardware and software configurations to the environments in which they operate. However, the use of MDPs in embedded signal processing systems is limited because of the large computational demands for solving this class of system models. This paper presents Sparse Parallel Value Iteration (SPVI), a new algorithm for solving large MDPs on resource-constrained embedded systems that are equipped with mobile GPUs. SPVI leverages recent advances in parallel solving of MDPs and adds sparse linear algebra techniques to significantly outperform the state-of-the-art. The method and its application are described in detail, and demonstrated with case studies that are implemented on an NVIDIA Tegra K1 System On Chip (SoC). The experimental results show execution time improvements in the range of 65 % -78% for several applications. SPVI also lifts restrictions required by other MDP solver approaches, making it more widely compatible with large classes of optimization problems.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Research area: Computer engineering, Computing Sciences, University of Maryland, Department of Electrical and Computer Engineering, Georgia Institute of Technology

Contributors: Sapio, A., Bhattacharyya, S. S., Wolf, M.

Number of pages: 6

Pages: 13-18

Publication date: Dec 2018

Host publication information

Title of host publication: 2018 Conference on Design and Architectures for Signal and Image Processing, DASIP 2018

Publisher: IEEE COMPUTER SOCIETY PRESS

ISBN (Electronic): 9781538682371

Publication series

Name: Conference on Design and Architectures for Signal and Image Processing, DASIP

ISSN (Print): 2164-9766

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Hardware and Architecture, Signal Processing, Electrical and Electronic Engineering

Keywords: CUDA, GPU, Markov decision processes, MDP, Sparsity, Value iteration

DOIs:

10.1109/DASIP.2018.8596969

Bibliographical note

jufoid=71852

Source: Scopus

Source ID: 85061388518

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

BM3D image denoising using heterogeneous computing platforms

Noise reduction is often performed at an early stage of the image processing path. In order to keep the processing delays small in different computing platforms, it is important that the noise reduction is performed swiftly. In this paper, the block-matching and three-dimensional filtering (BM3D) denoising algorithm is implemented on heterogeneous computing platforms using OpenCL and CUDA frameworks. To our knowledge, these implementations are the first successful open source attempts to use GPU computation for BM3D denoising. The presented GPU implementations are up to 7.5 times faster than their respective CPU implementations. At the same time, the experiments illustrate general design challenges in using massively parallel processing platforms for the calculation of complex imaging algorithms.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing Research Community (SPRC), Univ of Oulu, Center for Machine Vision Research

Contributors: Sarjanoja, S., Boutellier, J., Hannuksela, J.

Publication date: 28 Dec 2015

Host publication information

Title of host publication: DASIP 2015 - Proceedings of the 2015 Conference on Design and Architectures for Signal and Image Processing

Volume: 2015-December

Publisher: IEEE COMPUTER SOCIETY PRESS

Article number: 7367257

ISBN (Electronic): 9791092279108

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition, Hardware and Architecture, Signal Processing, Electrical and Electronic Engineering

Keywords: Image denoising, Mobile computing, Parallel algorithms, Parallel processing

DOIs:

10.1109/DASIP.2015.7367257

URLs:

<http://www.scopus.com/inward/record.url?scp=84959887479&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84959887479

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Reading on-screen text with gaze-based auto-scrolling

Visual information on eye movements can be used to facilitate scrolling while one is reading on-screen text. We carried out an experiment to find preferred reading regions on the screen and implemented an automatic scrolling technique based on the preferred regions of each individual reader. We then examined whether manual and automatic scrolling have an effect on reading behaviour on the basis of eye movement metrics, such as fixation duration and fixation count. We also studied how different font sizes affect the eye movement metrics. Results of analysis of data collected from 24 participants indicated no significant difference between manual and automatic scrolling in reading behaviour. Preferred reading regions on the screen varied among the participants. Most of them preferred relatively short regions. A significant effect of font size on fixation count was found. Subjective opinions indicated that participants found automatic scrolling convenient to use.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA)

Contributors: Sharmin, S., Špakov, O., Rähkä, K. J.

Number of pages: 8
Pages: 24-31
Publication date: 2013

Host publication information

Title of host publication: Proceedings of the 2013 Conference on Eye Tracking South Africa, ETSA 2013
ISBN (Print): 9781450321105
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: automatic scrolling, eye movements, fixation count, fixation duration, manual scrolling, reading, reading region
DOIs:
10.1145/2509315.2509319
URLs:
<http://www.scopus.com/inward/record.url?scp=84883884057&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84883884057
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

The effect of region of interest size on textural parameters

Texture analysis provides quantitative information describing the properties of a digital image. The value of texture analysis has been tested in various medical applications, using mostly magnetic resonance images because of the amount of information the method is capable to provide. However, there exists no certain practice to define the region of interest (ROI) within the texture parameters are calculated. Many parameters seem to be dependent on the ROI size. We studied the effect of the ROI size with magnetic resonance head images from 64 healthy adults and artificial noise images. According to our results, ROI size has a significant effect on the computed value of several second-order texture features. We conclude that comparisons of different size ROIs will therefore lead to falsely optimistic classification between analyzed tissues.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Department of Electronics and Communications Engineering, Research group: Quantative medical imaging , Integrated Technologies for Tissue Engineering Research (ITTE), Tampere University Hospital, Department of Radiology
Contributors: Sikiö, M., Holli-Helenius, K. K., Ryymin, P., Dastidar, P., Eskola, H., Harrison, L.
Number of pages: 5
Pages: 149-153
Publication date: 23 Oct 2015

Host publication information

Title of host publication: 2015 9th International Symposium on Image and Signal Processing and Analysis (ISPA)
Publisher: IEEE
ISBN (Electronic): 9781467380324
ASJC Scopus subject areas: Signal Processing, Computer Vision and Pattern Recognition
Keywords: magnetic resonance imaging, random pattern, region of interest, size, texture analysis
DOIs:
10.1109/ISPA.2015.7306049

Bibliographical note

EXT="Dastidar, Prasun"
EXT="Sikiö, Minna"
Source: Scopus
Source ID: 84978524965
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Liking the game: How can spectating motivations influence social media usage at live esports events?

There is no doubt that various social media services shape the ways in which we approach our daily lives. The ubiquitous nature of these services, afforded by mobile devices, means that we can take them with us wherever we go — including when we attend live events. Uncovering why individuals use social media during live events can help improve event organization, marketing, and the experiences of attendees. Our understanding of the motivations for using social media during live events is, however, still lacking in depth, especially in regard to emerging live events such as esports. This study aims to answer the question: what motivates the use of social media during live esports events? Data was gathered via a survey (N=255) at the 'Assembly 2016' LAN-event, a major live esports event. We examine the relationships between using various social media services and the motivations for esports spectating, through the Motivation Scale for Sports Consumption. While the results indicate that using social media services while attending Assembly 2016 was quite popular, it seemed that in many cases social media usage was a distraction from esports spectating, a core activity of the

event. The results provide implications as to how marketers of live esports events should encourage or control usage of social media by attendees.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Pervasive Computing, Gamification Group, Gamification Group
Contributors: Sjöblom, M., Hassan, L., Macey, J., Törhönen, M., Hamari, J.
Number of pages: 8
Pages: 160-167
Publication date: 18 Jul 2018

Host publication information

Title of host publication: Proceedings of the 9th International Conference on Social Media and Society, SMSociety 2018
Publisher: ACM
ISBN (Print): 9781450363341
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Esports, Live events, Motivation, Social media, Sports consumption
Electronic versions:
ACCEPTED_VERSION_Liking_the_game
DOIs:
10.1145/3217804.3217908
URLs:
<http://urn.fi/URN:NBN:fi:tuni-201912237103>

Bibliographical note

EXT="Törhönen, Maria"
DUPL=44481582
Source: Scopus
Source ID: 85051509297
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Interaction strategies for an affective conversational agent

The development of embodied conversational agents (ECA) as companions brings several challenges for both affective and conversational dialogue. These include challenges in generating appropriate affective responses, selecting the overall shape of the dialogue, providing prompt system response times, and handling interruptions. We present an implementation of such a companion showing the development of individual modules that attempt to address these challenges. Further, to resolve resulting conflicts, we present encompassing interaction strategies that attempt to balance the competing requirements along with dialogues from our working prototype to illustrate these interaction strategies in operation. Finally, we provide the results of an evaluation of the companion using an evaluation methodology created for conversational dialogue and including analysis using appropriateness annotation.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Augmented Human Activities (AHA), School of Computing Teesside University Middlesbrough, University of Oxford, Telefonica, School of Management (JKK), School of Computing Edinburgh Napier University Edinburgh, SICS SE-164 29 Kista, ILS Institute SUNY Albany Albany
Contributors: Smith, C., Crook, N., Dobnik, S., Charlton, D., Boye, J., Pulman, S., Santos de la Camara, R., Turunen, M., Benyon, D., Bradley, J., Gambäck, B., Hansen, P., Mival, O., Webb, N., Cavazza, M.
Number of pages: 17
Pages: 395-411
Publication date: Oct 2011
Peer-reviewed: Yes

Publication information

Journal: Presence: Teleoperators and Virtual Environments
Volume: 20
Issue number: 5
ISSN (Print): 1054-7460
Ratings:
Scopus rating (2011): CiteScore 3 SJR 0.354 SNIP 1.141
Original language: English
ASJC Scopus subject areas: Control and Systems Engineering, Software, Human-Computer Interaction, Computer Vision and Pattern Recognition

DOIs:

10.1162/PRES_a_00063

URLs:

<http://www.scopus.com/inward/record.url?scp=84863122938&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84863122938

Research output: [Contribution to journal](#) › [Article](#) › [Scientific](#) › [peer-review](#)

Multimodal subspace support vector data description

In this paper, we propose a novel method for projecting data from multiple modalities to a new subspace optimized for one-class classification. The proposed method iteratively transforms the data from the original feature space of each modality to a new common feature space along with finding a joint compact description of data coming from all the modalities. For data in each modality, we define a separate transformation to map the data from the corresponding feature space to the new optimized subspace by exploiting the available information from the class of interest only. We also propose different regularization strategies for the proposed method and provide both linear and non-linear formulations. The proposed Multimodal Subspace Support Vector Data Description outperforms all the competing methods using data from a single modality or fusing data from all modalities in four out of five datasets.

General information

Publication status: E-pub ahead of print

MoE publication type: A1 Journal article-refereed

Organisations: Computing Sciences, Research group: Multimedia Research Group - MRG, Finnish Environment Institute, Aarhus Universitet

Contributors: Sohrab, F., Raitoharju, J., Iosifidis, A., Gabbouj, M.

Number of pages: 13

Publication date: 2020

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition

Volume: 110

Article number: 107648

ISSN (Print): 0031-3203

Original language: English

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Artificial Intelligence

Keywords: Feature transformation, Multimodal data, One-class classification, Subspace learning, Support vector data description

Electronic versions:

Multimodal subspace support vector 2020

DOIs:

10.1016/j.patcog.2020.107648

URLs:

<http://urn.fi/URN:NBN:fi:tuni-202009297162>

Bibliographical note

EXT="Iosifidis, Alexandros"

Source: Scopus

Source ID: 85090954677

Research output: [Contribution to journal](#) › [Article](#) › [Scientific](#) › [peer-review](#)

PIVO: Probabilistic inertial-visual odometry for occlusion-robust navigation

This paper presents a novel method for visual-inertial odometry. The method is based on an information fusion framework employing low-cost IMU sensors and the monocular camera in a standard smartphone. We formulate a sequential inference scheme, where the IMU drives the dynamical model and the camera frames are used in coupling trailing sequences of augmented poses. The novelty in the model is in taking into account all the cross-terms in the updates, thus propagating the inter-connected uncertainties throughout the model. Stronger coupling between the inertial and visual data sources leads to robustness against occlusion and feature-poor environments. We demonstrate results on data collected with an iPhone and provide comparisons against the Tango device and using the EuRoC data set.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Aalto University

Contributors: Solin, A., Cortés, S., Rahtu, E., Kannala, J.

Number of pages: 10
Pages: 616-625
Publication date: 3 May 2018

Host publication information

Title of host publication: Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018
Publisher: IEEE
ISBN (Electronic): 9781538648865
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Science Applications
DOIs:
10.1109/WACV.2018.00073
Source: Scopus
Source ID: 85050916529
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Inertial Odometry on Handheld Smartphones

Building a complete inertial navigation system using the limited quality data provided by current smartphones has been regarded challenging, if not impossible. This paper shows that by careful crafting and accounting for the weak information in the sensor samples, smartphones are capable of pure inertial navigation. We present a probabilistic approach for orientation and use-case free inertial odometry, which is based on double-integrating rotated accelerations. The strength of the model is in learning additive and multiplicative IMU biases online. We are able to track the phone position, velocity, and pose in realtime and in a computationally lightweight fashion by solving the inference with an extended Kalman filter. The information fusion is completed with zero-velocity updates (if the phone remains stationary), altitude correction from barometric pressure readings (if available), and pseudo-updates constraining the momentary speed. We demonstrate our approach using an iPad and iPhone in several indoor dead-reckoning applications and in a measurement tool setup.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Signal Processing, Research group: Artificial Intelligence and Vision - AIV, Aalto University
Contributors: Solin, A., Cortes, S., Rahtu, E., Kannala, J.
Number of pages: 8
Pages: 1361-1368
Publication date: 5 Sep 2018

Host publication information

Title of host publication: 2018 21st International Conference on Information Fusion, FUSION 2018
Publisher: IEEE
Article number: 8455482
ISBN (Print): 9780996452762
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Signal Processing, Statistics, Probability and Uncertainty, Instrumentation
DOIs:
10.23919/ICIF.2018.8455482
Source: Scopus
Source ID: 85054102788
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Image-based characterization of the pulp flows

Material flow characterization is important in the process industries and its further automation. In this study, close-to-laminar pulp suspension flows are analyzed based on double-exposure images captured in laboratory conditions. The correlation-based methods including autocorrelation and the particle image pattern technique were studied. During the experiments, synthetic and real test data with manual ground truth were used. The particle image pattern matching method showed better performance achieving the accuracy of 90.0% for the real data set with linear motion of the suspension and 79.2% for the data set with flow distortions.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Department of Signal Processing, Lappeenranta University of Technology, Machine Vision and Pattern Recognition Laboratory, Laboratory of Biosystem Dynamics, Univ of Oulu, Monash University Malaysia
Contributors: Sorokin, M., Strokina, N., Eerola, T., Lensu, L., Karttunen, K., Kalviainen, H.
Number of pages: 8
Pages: 630-637

Publication date: 1 Jul 2016

Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition and Image Analysis

Volume: 26

Issue number: 3

ISSN (Print): 1054-6618

Ratings:

Scopus rating (2016): CiteScore 0.7 SJR 0.255 SNIP 0.872

Original language: English

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Vision and Pattern Recognition

Keywords: double-exposure, particle image velocimetry, pulp flow estimation

DOIs:

10.1134/S1054661816030196

Source: Scopus

Source ID: 84984924424

Research output: Contribution to journal › Article › Scientific › peer-review

Real-time hidden gaze point correction

The accuracy of gaze point estimation is one of the main limiting factors in developing applications that utilize gaze input. The existing gaze point correction methods either do not support real-time interaction or imply restrictions on gazecontrolled tasks and object screen locations. We hypothesize that when gaze points can be reliably correlated with object screen locations, it is possible to gather and leverage this information for improving the accuracy of gaze pointing. We propose an algorithm that uses a growing pool of such collected correlations between gaze points and objects for real-time hidden gaze point correction. We tested this algorithm assuming that any point inside of a rectangular object has equal probability to be hit by gaze. We collected real data in a user study to simulate pointing at targets of small (80px) size. The results showed that our algorithm can significantly improve the hit rate especially in pointing at middle-sized targets. The proposed method is real-time, person- and taskindependent and is applicable for arbitrary located objects.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Augmented Human Activities (AHA)

Contributors: Špakov, O., Gizatdinova, Y.

Number of pages: 4

Pages: 291-294

Publication date: 2014

Host publication information

Title of host publication: Proceedings of the Symposium on Eye Tracking Research and Applications, ETRA 2014

Publisher: Association for Computing Machinery

ISBN (Print): 9781450327510

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems

Keywords: Accuracy correction, Algorithms, Cumulative distribution function, Eye tracking, Gaze point, Pointing

DOIs:

10.1145/2578153.2578200

URLs:

<http://www.scopus.com/inward/record.url?scp=84899672400&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84899672400

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Look and lean: Accurate head-assisted eye pointing

Compared to the mouse, eye pointing is inaccurate. As a consequence, small objects are difficult to point by gaze alone. We suggest using a combination of eye pointing and subtle head movements to achieve accurate hands-free pointing in a conventional desktop computing environment. For tracking the head movements, we exploited information of the eye position in the eye tracker's camera view. We conducted a series of three experiments to study the potential caveats and benefits of using head movements to adjust gaze cursor position. Results showed that head-assisted eye pointing significantly improves the pointing accuracy without a negative impact on the pointing time. In some cases participants were able to point almost 3 times closer to the target's center, compared to the eye pointing alone (7 vs. 19 pixels). We conclude that head assisted eye pointing is a comfortable and potentially very efficient alternative for other assisting methods in the eye pointing, such as zooming.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Augmented Human Activities (AHA)
Contributors: Špakov, O., Isokoski, P., Majaranta, P.
Number of pages: 8
Pages: 35-42
Publication date: 2014

Host publication information

Title of host publication: Proceedings of the Symposium on Eye Tracking Research and Applications, ETRA 2014
Publisher: Association for Computing Machinery
ISBN (Print): 9781450327510
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems
Keywords: Eye tracking, Gaze input, Head movements, Pointing
DOIs:
10.1145/2578153.2578157
URLs:
<http://www.scopus.com/inward/record.url?scp=84899691537&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84899691537
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Comparison of eye movement filters used in HCI

We compared various real-time filters designed to denoise eye movements from low-sampling devices. Most of the filters found in literature were implemented and tested on data gathered in a previous study. An improvement was proposed for one of the filters. Parameters of each filter were adjusted to ensure their best performance. Four estimation parameters were proposed as criteria for comparison. The output from the filters was compared against two idealized signals (the signals denoised offline). The study revealed that FIR filters with triangular or Gaussian kernel (weighting) functions and parameters dependent on signal state show the best performance.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Augmented Human Activities (AHA)
Contributors: Špakov, O.
Number of pages: 4
Pages: 281-284
Publication date: 2012

Host publication information

Title of host publication: Proceedings - ETRA 2012: Eye Tracking Research and Applications Symposium
ISBN (Print): 9781450312257
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems
Keywords: algorithms, eye tracking, filters, gaze, smoothing
DOIs:
10.1145/2168556.2168616
URLs:
<http://www.scopus.com/inward/record.url?scp=84862667279&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 84862667279
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Comparison of gaze-to-objects mapping algorithms

Gaze data processing is an important and necessary step in gaze-based applications. This study focuses on the comparison of several gaze-to-object mapping algorithms using various dwell times for selection and presenting targets of several types and sizes. Seven algorithms found in literature were compared against two newly designed algorithms. The study revealed that a fractional mapping algorithm (known) has produced the highest rate of correct selections and fastest selection times, but also the highest rate of incorrect selections. The dynamic competing algorithm (designed) has shown the next best result, but also high rate of incorrect selections. A small impact on the type of target to the calculated statistics has been observed. A strictly centered gazing has helped to increase the rate of correct selections for all algorithms and types of targets. The directions for further mapping algorithms improvement and future investigation have been explained.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Augmented Human Activities (AHA)
Contributors: Špakov, O.
Publication date: 2011

Host publication information

Title of host publication: Proceedings of the 1st Conference on Novel Gaze-Controlled Applications, NGCA'11
Article number: 6
ISBN (Print): 9781450306805
ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software
Keywords: Algorithm design, Eye gaze pointing and selection, Gaze controlled applications, Gaze to object mapping
DOIs:
10.1145/1983302.1983308
URLs:
<http://www.scopus.com/inward/record.url?scp=79960161638&partnerID=8YFLogxK> (Link to publication in Scopus)
Source: Scopus
Source ID: 79960161638
Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Detection of bubbles as concentric circular arrangements

The paper proposes a method for the detection of bubble-like transparent objects in a liquid. The detection problem is non-trivial since bubble appearance varies considerably due to different lighting conditions causing contrast reversal and multiple interreflections. We formulate the problem as the detection of concentric circular arrangements (CCA). The CCAs are recovered in a hypothesize-optimize-verify framework. The hypothesis generation is based on sampling from the partially linked components of the non-maximum suppressed responses of oriented ridge filters, and is followed by the CCA parameter estimation. Parameter optimization is carried out by minimizing a novel cost-function. The performance was tested on gas dispersion images of pulp suspension and oil dispersion images. The mean error of gas/oil volume estimation was used as a performance criterion due to the fact that the main goal of the applications driving the research was the bubble volume estimation. The method achieved 28 and 13 % of gas and oil volume estimation errors correspondingly outperforming the OpenCV Circular Hough Transform in both cases and the WaldBoost detector in gas volume estimation.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Department of Signal Processing, Research group: Vision, Research Community on Data-to-Decision (D2D), Machine Vision and Pattern Recognition Laboratory, Lappeenranta University of Technology, Computer Vision Group, Czech Technical University in Prague, Monash University Malaysia
Contributors: Strokina, N., Matas, J., Eerola, T., Lensu, L., Kälviäinen, H.
Number of pages: 10
Pages: 387-396
Publication date: Apr 2016
Peer-reviewed: Yes
Early online date: 10 Feb 2016

Publication information

Journal: Machine Vision and Applications
Volume: 27
Issue number: 3
ISSN (Print): 0932-8092
Ratings:
Scopus rating (2016): CiteScore 4.7 SJR 0.741 SNIP 1.433
Original language: English
ASJC Scopus subject areas: Hardware and Architecture, Computer Vision and Pattern Recognition, Software, Computer Science Applications
Keywords: Bubble detection, Circular arrangements, Image processing, Machine vision, Object segmentation
DOIs:
10.1007/s00138-016-0749-7
Source: Scopus
Source ID: 84957656160
Research output: Contribution to journal › Article › Scientific › peer-review

Microservices in agile software development: A workshop-based study into issues, advantages, and disadvantages

In the last years, cloud-native architectures have emerged as a target platform for the deployment of microservice architectures. The migration of existing monoliths into cloud-native applications is still in the early phase, and only few companies already started their migrations. Therefore, success and failure stories about different approaches are not available in the literature. This context connects also to the recently discussed DevOps context where development and continuous deployment are closely linked.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Free University of Bolzano-Bozen, Former organisation of the author

Contributors: Taibi, D., Lenarduzzi, V., Pahl, C., Janes, A.

Publication date: 22 May 2017

Host publication information

Title of host publication: Proceedings of the XP2017 Scientific Workshops, XP 2017

Volume: Part F129907

Publisher: Association for Computing Machinery

Article number: a23

ISBN (Electronic): 9781450352642

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Cloud software, Microservices, SOA, Software architecture

DOIs:

10.1145/3120459.3120483

Source: Scopus

Source ID: 85029863670

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Operationalizing the experience factory for effort estimation in agile processes

[Background] The effort required to systematically collect historical data is not always allocable in agile processes and historical data management is usually delegated to the developers' experience, who need to remember previous project details. However, even if well trained, developers cannot precisely remember a huge number of details, resulting in wrong decisions being made during the development process. [Aims] The goal of this paper is to operationalize the Experience Factory in an agile way, i.e., defining a strategy for collecting historical project data using an agile approach. [Method] We provide a mechanism for understanding whether a measure must be collected or not, based on the Return on Invested Time (ROIT). In order to validate this approach, we instantiated the factory with an exploratory case study, comparing four projects that did not use our approach with one project that used it after 12 weeks out of 37 and two projects that used it from the beginning. [Results] The proposed approach helps developers to constantly improve their estimation accuracy with a very positive ROIT of the collected measure. [Conclusions] From this first experience, we can conclude that the Experience Factory can be applied effectively to agile processes, supporting developers in improving their performance and reducing potential decision mistakes.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Fraunhofer IESE, University of Cagliari, Former organisation of the author

Contributors: Taibi, D., Lenarduzzi, V., Diebold, P., Lunesu, I.

Number of pages: 10

Pages: 31-40

Publication date: 15 Jun 2017

Host publication information

Title of host publication: Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering, EASE 2017

Volume: Part F128635

Publisher: Association for Computing Machinery

ISBN (Electronic): 9781450348041

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Agile software development, Experience factory, Knowledge management

DOIs:

10.1145/3084226.3084240

Source: Scopus

Source ID: 85025449243

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Comparing communication effort within the scrum, scrum with Kanban, XP, and Banana development processes

[Context]: Communication plays an important role in any development process. However, communication overhead has been rarely compared among development processes. [Objective]: The goal of this work is to compare the communication overhead and the different channels applied in three agile processes (XP, Scrum, Scrum with Kanban) and in an unstructured process. [Method]: We designed an empirical study asking four teams to develop the same application with the four development processes, and we compare the communication overhead among them. [Results]: As expected, face-to-face communication is most frequently employed in the teams. Scrum with Kanban turned out to be the process that requires the least communication. Unexpectedly, despite requiring much more time to develop the same application, the unstructured process required comparable communication overhead (25% of the total development time) as the agile processes.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: University of Oulu, Former organisation of the author

Contributors: Taibi, D., Lenarduzzi, V., Ahmad, M. O., Liukkunen, K.

Number of pages: 6

Pages: 258-263

Publication date: 15 Jun 2017

Host publication information

Title of host publication: Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering, EASE 2017

Volume: Part F128635

Publisher: Association for Computing Machinery

ISBN (Electronic): 9781450348041

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Agile processes, Case study, Communication, Empirical software engineering

DOIs:

10.1145/3084226.3084270

Source: Scopus

Source ID: 85025468824

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

The Accuracy of Atrial Fibrillation Detection from Wrist Photoplethysmography. A Study on Post-Operative Patients

Atrial fibrillation (AF) is the most common type of cardiac arrhythmia. Although not life-threatening itself, AF significantly increases the risk of stroke and myocardial infarction. Current tools available for screening and monitoring of AF are inadequate and an unobtrusive alternative, suitable for long-term use, is needed. This paper evaluates an atrial fibrillation detection algorithm based on wrist photoplethysmographic (PPG) signals. 29 patients recovering from surgery in the post-anesthesia care unit were monitored. 15 patients had sinus rhythm (SR, 67.5± 10.7 years old, 7 female) and 14 patients had AF (74.8± 8.3 years old, 8 female) during the recordings. Inter-beat intervals (IBI) were estimated from PPG signals. As IBI estimation is highly sensitive to motion or other types of noise, acceleration signals and PPG waveforms were used to automatically detect and discard unreliable IBI. AF was detected from windows of 20 consecutive IBI with 98.45±6.89% sensitivity and 99.13±1.79% specificity for 76.34±19.54% of the time. For the remaining time, no decision was taken due to the lack of reliable IBI. The results show that wrist PPG is suitable for long term monitoring and AF screening. In addition, this technique provides a more comfortable alternative to ECG devices.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Faculty of Biomedical Sciences and Engineering, Research group: Sensor Technology and Biomeasurements (STB), PulseOn SA, Tampere University Hospital

Contributors: Tarniceriu, A., Harju, J., Yousefi, Z. R., Vehkaoja, A., Parak, J., Yli-Hankala, A., Korhonen, I.

Number of pages: 4

Pages: 4844-4847

Publication date: 26 Oct 2018

Host publication information

Title of host publication: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2018

Volume: 2018-July

Publisher: IEEE

Article number: 8513197

ISBN (Electronic): 9781538636466

ASJC Scopus subject areas: Signal Processing, Biomedical Engineering, Computer Vision and Pattern Recognition, Health Informatics

Electronic versions:

The Accuracy of Atrial Fibrillation Detection from Wrist Photoplethysmography. A study on Post-Operative Patients
DOIs:

10.1109/EMBC.2018.8513197

URLs:

<http://urn.fi/URN:NBN:fi:tty-201902051214>

Bibliographical note

INT=tut-bmt, "Yousefi, Zeinab Rezaei"

Source: Scopus

Source ID: 85056672654

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Digging deeper into egocentric gaze prediction

This paper digs deeper into factors that influence egocentric gaze. Instead of training deep models for this purpose in a blind manner, we propose to inspect factors that contribute to gaze guidance during daily tasks. Bottom-up saliency and optical flow are assessed versus strong spatial prior baselines. Task-specific cues such as vanishing point, manipulation point, and hand regions are analyzed as representatives of top-down information. We also look into the contribution of these factors by investigating a simple recurrent neural model for ego-centric gaze prediction. First, deep features are extracted for all input video frames. Then, a gated recurrent unit is employed to integrate information over time and to predict the next fixation. We also propose an integrated model that combines the recurrent model with several top-down and bottom-up cues. Extensive experiments over multiple datasets reveal that (1) spatial biases are strong in egocentric videos, (2) bottom-up saliency models perform poorly in predicting gaze and underperform spatial biases, (3) deep features perform better compared to traditional features, (4) as opposed to hand regions, the manipulation point is a strong influential cue for gaze prediction, (5) combining the proposed recurrent model with bottom-up cues, vanishing points and, in particular, manipulation point results in the best gaze prediction accuracy over egocentric videos, (6) the knowledge transfer works best for cases where the tasks or sequences are similar, and (7) task and activity recognition can benefit from gaze prediction. Our findings suggest that (1) there should be more emphasis on hand-object interaction and (2) the egocentric vision community should consider larger datasets including diverse stimuli and more subjects.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Aalto University

Contributors: Tavakoli, H. R., Rahtu, E., Kannala, J., Borji, A.

Number of pages: 10

Pages: 273-282

Publication date: 4 Mar 2019

Host publication information

Title of host publication: 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019

Publisher: IEEE

ISBN (Electronic): 9781728119755

Publication series

Name: IEEE Winter Conference on Applications of Computer Vision

ISSN (Print): 1550-5790

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Science Applications

DOIs:

10.1109/WACV.2019.00035

Bibliographical note

jufoid=57596

Source: Scopus

Source ID: 85063594608

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Deep audio-visual saliency: Baseline model and data

This paper introduces a conceptually simple and effective Deep Audio-Visual Embedding for dynamic saliency prediction dubbed "DAVE" in conjunction with our efforts towards building an Audio-Visual Eye-tracking corpus named "AVE". Despite existing a strong relation between auditory and visual cues for guiding gaze during perception, video saliency models only consider visual cues and neglect the auditory information that is ubiquitous in dynamic scenes. Here, we propose a baseline deep audio-visual saliency model for multi-modal saliency prediction in the wild. Thus the proposed model is intentionally designed to be simple. A video baseline model is also developed on the same architecture to assess effectiveness of the audio-visual models on a fair basis. We demonstrate that audio-visual saliency model outperforms the video saliency models. The data and code are available at <https://hrtavakoli.github.io/AVE/> and <https://github.com/hrtavakoli/DAVE>.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Computing Sciences, Nokia, Aalto University
Contributors: Tavakoli, H. R., Borji, A., Kannala, J., Rahtu, E.
Publication date: 6 Feb 2020

Host publication information

Title of host publication: Proceedings ETRA 2020 Short Papers - ACM Symposium on Eye Tracking Research and Applications, ETRA 2020
Publisher: ACM
Editor: Spencer, S. N.
Article number: 3
ISBN (Electronic): 9781450371346
ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Human-Computer Interaction, Ophthalmology, Sensory Systems
Keywords: Audio-Visual Saliency, Deep Learning, Dynamic Visual Attention
DOIs:
10.1145/3379156.3391337

Bibliographical note

EXT="Tavakoli, Hamed Rezazadegan"
Source: Scopus
Source ID: 85085734752
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Multilinear class-specific discriminant analysis

There has been a great effort to transfer linear discriminant techniques that operate on vector data to high-order data, generally referred to as Multilinear Discriminant Analysis (MDA) techniques. Many existing works focus on maximizing the inter-class variances to intra-class variances defined on tensor data representations. However, there has not been any attempt to employ class-specific discrimination criteria for the tensor data. In this paper, we propose a multilinear subspace learning technique suitable for applications requiring class-specific tensor models. The method maximizes the discrimination of each individual class in the feature space while retains the spatial structure of the input. We evaluate the efficiency of the proposed method on two problems, i.e. facial image analysis and stock price prediction based on limit order book data.

General information

Publication status: Published
MoE publication type: A1 Journal article-refereed
Organisations: Signal Processing, Aarhus Universitet, Laboratory of Signal Processing
Contributors: Thanh Tran, D., Gabbouj, M., Iosifidis, A.
Number of pages: 6
Pages: 131-136
Publication date: 1 Dec 2017
Peer-reviewed: Yes

Publication information

Journal: Pattern Recognition Letters
Volume: 100
ISSN (Print): 0167-8655
Ratings:
Scopus rating (2017): CiteScore 5.5 SJR 0.662 SNIP 1.605
Original language: English
ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Artificial Intelligence

Keywords: Class-specific discriminant learning, Face verification, Multilinear discriminant analysis, Stock price prediction
DOIs:

10.1016/j.patrec.2017.10.027

URLs:

<https://arxiv.org/abs/1710.10695>

Bibliographical note

INT=sgn,"Thanh Tran, Dat"

Source: Scopus

Source ID: 85032300703

Research output: Contribution to journal > Article > Scientific > peer-review

Eigen Posture Based Fall Risk Assessment System Using Kinect

Postural Instability (PI) is a major reason for fall in geriatric population as well as for people with diseases or disorders like Parkinson's, stroke etc. Conventional stability indicators like Berg Balance Scale (BBS) require clinical settings with skilled personnel's interventions to detect PI and finally classify the person into low, mid or high fall risk categories. Moreover these tests demand a number of functional tasks to be performed by the patient for proper assessment. In this paper a machine learning based approach is developed to determine fall risk with minimal human intervention using only Single Limb Stance exercise. The analysis is done based on the spatiotemporal dynamics of skeleton joint positions obtained from Kinect sensor. A novel posture modeling method has been applied for feature extraction along with some traditional time domain and metadata features to successfully predict the fall risk category. The proposed unobtrusive, affordable system is tested over 224 subjects and is able to achieve 75% mean accuracy on the geriatric and patient population.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Tata Consultancy Services India

Contributors: Tripathy, S. R., Chakravarty, K., Sinha, A.

Number of pages: 4

Pages: 1-4

Publication date: 26 Oct 2018

Host publication information

Title of host publication: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2018

Volume: 2018-July

Publisher: IEEE

Article number: 8513263

ISBN (Electronic): 9781538636466

ASJC Scopus subject areas: Signal Processing, Biomedical Engineering, Computer Vision and Pattern Recognition, Health Informatics

Keywords: BBS, Eigenpose, EMD, Fall risk, Index Terms-Kinect

DOIs:

10.1109/EMBC.2018.8513263

Source: Scopus

Source ID: 85056666030

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

ICface: Interpretable and controllable face reenactment using GANs

This paper presents a generic face animator that is able to control the pose and expressions of a given face image. The animation is driven by human interpretable control signals consisting of head pose angles and the Action Unit (AU) values. The control information can be obtained from multiple sources including external driving videos and manual controls. Due to the interpretable nature of the driving signal, one can easily mix the information between multiple sources (e.g. pose from one image and expression from another) and apply selective postproduction editing. The proposed face animator is implemented as a two stage neural network model that is learned in self-supervised manner using a large video collection. The proposed Interpretable and Controllable face reenactment network (ICface) is compared to the state-of-the-art neural network based face animation techniques in multiple tasks. The results indicate that ICface produces better visual quality, while being more versatile than most of the comparison methods. The introduced model could provide a lightweight and easy to use tool for multitude of advanced image and video editing tasks. The program code will be publicly available upon the acceptance of the paper.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Aalto University
Contributors: Tripathy, S., Kannala, J., Rahtu, E.
Number of pages: 10
Pages: 3374-3383
Publication date: 1 Mar 2020

Host publication information

Title of host publication: 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020
Publisher: IEEE
ISBN (Electronic): 9781728165530

Publication series

Name: IEEE Winter Conference on Applications of Computer Vision
ISSN (Print): 1550-5790
ASJC Scopus subject areas: Computer Science Applications, Computer Vision and Pattern Recognition
DOIs:
10.1109/WACV45572.2020.9093474

Bibliographical note

jufoid=57596
Source: Scopus
Source ID: 85085467341
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Dual Structured Convolutional Neural Network with Feature Augmentation for Quantitative Characterization of Tissue Histology

We present a dual convolutional neural network (dCNN) architecture for extracting multi-scale features from histological tissue images for the purpose of automated characterization of tissue in digital pathology. The dual structure consists of two identical convolutional neural networks applied to input images with different scales, that are merged together and stacked with two fully connected layers. It has been acknowledged that deep networks can be used to extract higher-order features, and therefore, the network output at final fully connected layer was used as a deep dCNN feature vector. Further, engineered features, shown in previous studies to capture important characteristics of tissue structure and morphology, were integrated to the feature extractor module. The acquired quantitative feature representation can be further utilized to train a discriminative model for classifying tissue types. Machine learning based methods for detection of regions of interest, or tissue type classification will advance the transition to decision support systems and computer aided diagnosis in digital pathology. Here we apply the proposed feature-augmented dCNN method with supervised learning in detecting cancerous tissue from whole slide images. The extracted quantitative representation of tissue histology was used to train a logistic regression model with elastic net regularization. The model was able to accurately discriminate cancerous tissue from normal tissue, resulting in blockwise AUC=0.97, where the total number of analyzed tissue blocks was approximately 8.3 million that constitute the test set of 75 whole slide images.

General information

Publication status: Published
MoE publication type: A4 Article in a conference publication
Organisations: Faculty of Biomedical Sciences and Engineering, Signal Processing
Contributors: Valkonen, M., Kartasalo, K., Liimatainen, K., Nykter, M., Latonen, L., Ruusuvoori, P.
Number of pages: 9
Pages: 27-35
Publication date: 19 Jan 2018

Host publication information

Title of host publication: 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017
Publisher: IEEE
ISBN (Electronic): 9781538610343
ASJC Scopus subject areas: Computer Science Applications, Computer Vision and Pattern Recognition
DOIs:
10.1109/ICCVW.2017.10

Bibliographical note

EXT="Nykter, Matti"
Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Guidelines for development and evaluation of usage data analytics tools for human-machine interactions with industrial manufacturing systems

We present the lessons learned during the development and evaluation process for UX-sensors, a visual data analytics tool for inspecting logged usage data from flexible manufacturing systems (FMS). Based on the experiences during a

collaborative development process with practitioners from one FMS supplier company, we propose guidelines to support other developers of visual data analytics tools for usage data logging in context of complex industrial systems. For instance, involving stakeholders with different roles can help to identify user requirements and generate valuable development ideas. Tool developers should confirm early access to real usage data from customers' systems and familiarize themselves with the log data structure. We argue that combining expert evaluations with field study methods can provide a more diverse set of usability issues to address. For future research, we encourage studies on insights emerging from usage data analytics and their impact on the viewpoints of the supplier and customer.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Pervasive Computing, Research area: User experience, University of Wisconsin-Stevens Point, Fastems

Contributors: Varsaluoma, J., Väättäjä, H., Heimonen, T., Tiitinen, K., Hakulinen, J., Turunen, M., Nieminen, H.

Number of pages: 10

Pages: 172-181

Publication date: 10 Oct 2018

Host publication information

Title of host publication: Mindtrek 2018 - Proceedings of the 22nd International Academic Mindtrek Conference

Publisher: ACM

ISBN (Electronic): 9781450365895

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

DOIs:

10.1145/3275116.3275138

Bibliographical note

EXT="Nieminen, Harri"

Source: Scopus

Source ID: 85056717713

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Sparse sampling for real-time ray tracing

Ray tracing is an interesting rendering technique, but remains too slow for real-time applications. There are various algorithmic methods to speed up ray tracing through uneven screen-space sampling, e.g., foveated rendering where sampling is directed by eye tracking. Uneven sampling methods tend to require at least one sample per pixel, limiting their use in real-time rendering. We review recent work on image reconstruction from arbitrarily distributed samples, and argue that these will play major role in the future of real-time ray tracing, allowing a larger fraction of samples to be focused on regions of interest. Potential implementation approaches and challenges are discussed.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Pervasive Computing, Research area: Computer engineering

Contributors: Viitanen, T., Koskela, M., Immonen, K., Mäkitalo, M., Jääskeläinen, P., Takala, J.

Number of pages: 8

Pages: 295-302

Publication date: 2018

Host publication information

Title of host publication: VISIGRAPP 2018 - Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications

Volume: 1

Publisher: SCITEPRESS

ISBN (Electronic): 9789897582875

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Graphics and Computer-Aided Design, Artificial Intelligence

Keywords: Image reconstruction, Ray tracing, Sparse sampling

DOIs:

10.5220/0006655802950302

Source: Scopus

Source ID: 85047764122

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Anthropometric clothing measurements from 3D body scans

We propose a full processing pipeline to acquire anthropometric measurements from 3D measurements. The first stage of our pipeline is a commercial point cloud scanner. In the second stage, a pre-defined body model is fitted to the captured point cloud. We have generated one male and one female model from the SMPL library. The fitting process is based on non-rigid iterative closest point algorithm that minimizes overall energy of point distance and local stiffness energy terms. In the third stage, we measure multiple circumference paths on the fitted model surface and use a nonlinear regressor to provide the final estimates of anthropometric measurements. We scanned 194 male and 181 female subjects, and the proposed pipeline provides mean absolute errors from 2.5 to 16.0 mm depending on the anthropometric measurement.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Computing Sciences, Research group: Vision, NOMO Technologies Ltd

Contributors: Yan, S., Wirta, J., Kämäräinen, J.

Number of pages: 11

Publication date: 2020

Peer-reviewed: Yes

Publication information

Journal: Machine Vision and Applications

Volume: 31

Issue number: 1-2

Article number: 7

ISSN (Print): 0932-8092

Original language: English

ASJC Scopus subject areas: Software, Hardware and Architecture, Computer Vision and Pattern Recognition, Computer Science Applications

Keywords: 3D body model, Anthropometric measurement, Non-rigid ICP

Electronic versions:

Anthropometric clothing measurements 2020

DOIs:

10.1007/s00138-019-01054-4

URLs:

<http://urn.fi/URN:NBN:fi:tuni-202006116044>

Source: Scopus

Source ID: 85078296322

Research output: Contribution to journal > Article > Scientific > peer-review

Hierarchical deformable part models for heads and tails

Imbalanced long-tail distributions of visual class examples inhibit accurate visual detection, which is addressed by a novel Hierarchical Deformable Part Model (HDPM). HDPM constructs a sub-category hierarchy by alternating bootstrapping and Visual Similarity Network (VSN) based discovery of head and tail sub-categories. We experimentally evaluate HDPM and compare with other sub-category aware visual detection methods with a moderate size dataset (Pascal VOC 2007), and demonstrate its scalability to a large scale dataset (ILSVRC 2014 Detection Task). The proposed HDPM consistently achieves significant performance improvement in both experiments.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Signal Processing, Research group: Vision

Contributors: Yancheshmeh, F. S., Chen, K., Kämäräinen, J.

Number of pages: 11

Pages: 45-55

Publication date: 2018

Host publication information

Title of host publication: VISIGRAPP 2018 - Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications

Volume: 5

Publisher: SCITEPRESS

ISBN (Electronic): 9789897582905

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Computer Graphics and Computer-Aided Design, Artificial Intelligence

Keywords: Deformable part model, Imbalanced datasets, Localization, Long-tail distribution, Object detection, Sub-category discovery, Visual similarity network

DOIs:

10.5220/0006532700450055

Source: Scopus

Source ID: 85047826548

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Visibility-Aware Part Coding for Vehicle Viewing Angle Estimation

A number of spatially-localised semantic parts of vehicles sensitive to pose changes are encoded their visible probabilities into a mid-level feature vector. Car pose estimation is then formulated into a regression on concatenated low-and mid-level features to continuously changing viewing angles. Each dimension of our visibility-Aware part codes separates all the training samples into two groups according to its visual existence in images, which provides additional part-specific range constraint of viewing angles. Moreover, the proposed codes can alleviate the suffering from sparse and imbalanced data distribution in the light of modelling latent dependency across angle targets. Experimental evaluation for car pose estimation on the EPFL Multi-View Car benchmark demonstrates significant improvement of our method over the state-of-The-Art regression methods, especially when only sparse and imbalanced data is available.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Research group: Vision, South China University of Technology

Contributors: Yang, D., Qian, Y., Cai, D., Yan, S., Kämäräinen, J., Chen, K.

Number of pages: 6

Pages: 65-70

Publication date: 1 Aug 2019

Host publication information

Title of host publication: 9th International Conference on Information Science and Technology, ICIST 2019

Publisher: IEEE

ISBN (Electronic): 9781728121062

ASJC Scopus subject areas: Computer Science Applications, Computer Vision and Pattern Recognition, Information Systems, Computational Mathematics, Control and Optimization

Keywords: Car pose estimation, Coding, Pose-sensitive parts, Regression forests, Visibility-Aware

DOIs:

10.1109/ICIST.2019.8836907

Bibliographical note

EXT="Chen, Ke"

jufoid=79229

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Looking for a five-legged sheep: Identifying enterprise architects' skills and competencies

Enterprise architecture (EA) is a holistic approach to comprehend the organization's business objectives and processes, data resources, information systems and information technologies. To advance EA activities, organizations need a myriad of different skills and competences both from individual enterprise architects and from architect teams. However, research on these skills and competences is scarce. Not knowing what skills are actually needed might be one of the reasons why public sector EA endeavors have been very problematic. In this paper, we conduct a qualitative survey among enterprise architects themselves to identify which skills they consider essential for EA work. Our results indicate that the range of skills is great, and finding an expert with all appropriate competencies is like looking for a fivelegged sheep.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Industrial and Information Management

Contributors: Ylinen, M., Pekkola, S.

Publication date: 30 May 2018

Host publication information

Title of host publication: Proceedings of the 19th Annual International Conference on Digital Government Research : Governance in the Data Age, DG.O 2018

Publisher: ACM

Article number: a58

ISBN (Electronic): 9781450365260

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

DOIs:

10.1145/3209281.3209353

Source: Scopus

Source ID: 85049050136

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Multi-view predictive latent space learning

In unsupervised circumstances, multi-view learning seeks a shared latent representation by taking the consensus and complementary principles into account. However, most existing multi-view unsupervised learning approaches do not explicitly lay stress on the predictability of the latent space. In this paper, we propose a novel multi-view predictive latent space learning (MVP) model and apply it to multi-view clustering and unsupervised dimension reduction. The latent space is forced to be predictive by maximizing the correlation between the latent space and feature space of each view. By learning a multi-view graph with adaptive view-weight learning, MVP effectively combines the complementary information from multi-view data. Experimental results on benchmark datasets show that MVP outperforms the state-of-the-art multi-view clustering and unsupervised dimension reduction algorithms.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Signal Processing, Tianjin University

Contributors: Yuan, J., Gao, K., Zhu, P., Egiazarian, K.

Publication date: 2018

Peer-reviewed: Yes

Early online date: 2018

Publication information

Journal: Pattern Recognition Letters

ISSN (Print): 0167-8655

Ratings:

Scopus rating (2018): CiteScore 5.8 SJR 0.662 SNIP 1.729

Original language: English

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Artificial Intelligence

Keywords: Multi-view learning, Predictive latent space learning, Unsupervised clustering, Unsupervised dimension reduction

DOIs:

10.1016/j.patrec.2018.06.022

Source: Scopus

Source ID: 85049094619

Research output: Contribution to journal › Article › Scientific › peer-review

Automated design of networks of transport-triggered architecture processors using dynamic dataflow programs

Modern embedded systems show a clear trend towards the use of Multiprocessor System-on-Chip (MPSoC) architectures in order to handle the performance and power consumption constraints. However, the design and validation of dedicated MPSoCs is an extremely hard and expensive task due to their complexity. Thus, the development of automated design processes is of highest importance to satisfy the time-to-market pressure of embedded systems. This paper proposes an automated co-design flow based on the high-level language-based approach of the Reconfigurable Video Coding framework. The designer provides the application description in the RVC-CAL dataflow language, after which the presented co-design flow automatically generates a network of heterogeneous processors that can be synthesized on FPGA chips. The synthesized processors are Very Long Instruction Word-style processors. Such a methodology permits the rapid design of a many-core signal processing system which can take advantage of all levels of parallelism. The toolchain functionality has been demonstrated by synthesizing an MPEG-4 Simple Profile video decoder to two different FPGA boards. The decoder is realized into 18 processors that decode QCIF resolution video at 45 frames per second on a 50 MHz FPGA clock frequency. The results show that the given application can take advantage of every level of parallelism.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Signal Processing Research Community (SPRC), Universite de Rennes, CSE Department, Univ of Oulu, UBL

Contributors: Yviquel, H., Boutellier, J., Raulet, M., Casseau, E.

Number of pages: 8

Pages: 1295-1302

Publication date: Nov 2013

Peer-reviewed: Yes

Publication information

Journal: Signal Processing: Image Communication

Volume: 28

Issue number: 10

ISSN (Print): 0923-5965

Ratings:

Scopus rating (2013): CiteScore 3.2 SJR 0.407 SNIP 1.301

Original language: English

ASJC Scopus subject areas: Software, Signal Processing, Computer Vision and Pattern Recognition, Electrical and Electronic Engineering

Keywords: Co-design, Dataflow programming, Multi-Processor System-on-Chip (MPSoC), Reconfigurable Video Coding (RVC), Transport-Trigger Architecture (TTA)

DOIs:

10.1016/j.image.2013.08.013

URLs:

<http://www.scopus.com/inward/record.url?scp=84888203042&partnerID=8YFLogxK> (Link to publication in Scopus)

Source: Scopus

Source ID: 84888203042

Research output: Contribution to journal › Article › Scientific › peer-review

Decoding complexity reduction in projection-based light-field 3D displays using self-contained HEVC tiles

The goal of this work is to provide a low complexity video decoding solution for High Efficiency Video Coding (HEVC) streams in applications where only a region of the video frames is needed to be decoded. This paper studies the problem of creating self-contained (i.e., independently decodable) partitions in the HEVC streams. Further, the requirements for building self-contained regions are described, and an encoder-side solution is proposed based on HEVC tile feature. A particular application of self-contained tiles targets the type of light-field 3D displays, which employ a dense set of optical engines to recreate the light field. Correspondingly, such 3D displays require a dense set of input views and therefore the partial decoding of bitstreams helps providing less complex and consequently real-time decoding and processing. The simulation results show a significant increase in decoding speed at the cost of a minor increase in storage capacity.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Department of Signal Processing, Research group: 3D MEDIA, Nokia

Contributors: Zare, A., Kovacs, P. T., Aminlou, A., Hannuksela, M. M., Gotchev, A.

Publication date: 22 Aug 2016

Host publication information

Title of host publication: 2016 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video, 3DTV-CON 2016

Publisher: IEEE COMPUTER SOCIETY PRESS

ISBN (Electronic): 9781509033133

ASJC Scopus subject areas: Computer Graphics and Computer-Aided Design, Computer Networks and Communications, Computer Vision and Pattern Recognition, Human-Computer Interaction, Electrical and Electronic Engineering

Keywords: HEVC, light-field 3D displays, partial decoding, random access, slice, tile, video partitioning

DOIs:

10.1109/3DTV.2016.7548965

Bibliographical note

INT=sgn,"Zare, Alireza"

Source: Scopus

Source ID: 84987792281

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review

Real-time online drilling vibration analysis using data mining

While the data mining intermediaries play a critical role in the rock drilling industry, they also tend to provide an optimized real-time model for the drilling systems. In addition, proper online tool condition monitoring (OTOM) methods can improve the drilling performance by accessing real-time data. Hence, OTOM methods assist depreciating error and detect unspecified faults at early stages. In this study, we proposed appropriate OTOM algorithms to develop and enhance the quality of real-time systems and provide a solution to detect and categorize various stages of drilling operation with the aid of vibration signals (especially in terms of acceleration or velocity). In particular, the proposed methods in this article perform based on statistical approaches. Therefore, in order to recognize the drilling stages, we measured the Root Mean Square (RMS) values corresponding to the acceleration signals. In the meantime, we also succeeded to distinguish the drilling stages by employing estimated power spectral density (PSD) in the frequency domain. The acquired results in this publication confirm the real-time prediction and classification potential of the proposed methods for the different drilling

stages and especially for the rock drilling engineering.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Automation Technology and Mechanical Engineering, Research group: Innovative Hydraulic Automation, Research group: MMDM, Tamlink Oy, Sandvik Mining and Construction Oy

Contributors: Zare, M., Huova, M., Visa, A., Launis, S.

Number of pages: 6

Pages: 175-180

Publication date: 19 Jul 2019

Host publication information

Title of host publication: Proceedings of the 2019 2nd International Conference on Data Science and Information Technology, DSIT 2019

Publisher: ACM

ISBN (Electronic): 9781450371414

ASJC Scopus subject areas: Human-Computer Interaction, Computer Networks and Communications, Computer Vision and Pattern Recognition, Software

Keywords: Data mining, Drilling stages, Real-time, Statistical analysis

DOIs:

10.1145/3352411.3352439

Source: Scopus

Source ID: 85072810540

Research output: Chapter in Book/Report/Conference proceeding > Conference contribution > Scientific > peer-review

Adaptive sampling for compressed sensing based image compression

The compressed sensing (CS) theory has been successfully applied to image compression in the past few years as most image signals are sparse in a certain domain. In this paper, we focus on how to improve the sampling efficiency for CS-based image compression by using our proposed adaptive sampling mechanism on the block-based CS (BCS), especially the reweighted one. To achieve this goal, two solutions are developed at the sampling side and reconstruction side, respectively. The proposed sampling mechanism allocates the CS-measurements to image blocks according to the statistical information of each block so as to sample the image more efficiently. A generic allocation algorithm is developed to help assign CS-measurements and several allocation factors derived in the transform domain are used to control the overall allocation in both solutions. Experimental results demonstrate that our adaptive sampling scheme offers a very significant quality improvement as compared with traditional non-adaptive ones.

General information

Publication status: Published

MoE publication type: A1 Journal article-refereed

Organisations: Department of Signal Processing, Research group: Video, Research Community on Data-to-Decision (D2D), Institute of Image Processing, University of Electronic Science and Technology of China, Hong Kong University of Science and Technology

Contributors: Zhu, S., Zeng, B., Gabbouj, M.

Number of pages: 12

Pages: 94-105

Publication date: 1 Jul 2015

Peer-reviewed: Yes

Publication information

Journal: Journal of Visual Communication and Image Representation

Volume: 30

ISSN (Print): 1047-3203

Ratings:

Scopus rating (2015): CiteScore 3.6 SJR 0.632 SNIP 1.525

Original language: English

ASJC Scopus subject areas: Electrical and Electronic Engineering, Media Technology, Computer Vision and Pattern Recognition, Signal Processing

Keywords: Adaptive sampling, Block-based compressed sensing (BCS), Image coding, Image compression, Measurement allocation, Sampling efficiency, Sparsity Compressed sensing (CS)

DOIs:

10.1016/j.jvcir.2015.03.006

Source: Scopus

Source ID: 84926618362

Research output: Contribution to journal > Article > Scientific > peer-review

End-to-end learning for video frame compression with self-attention

One of the core components of conventional (i.e., non-learned) video codecs consists of predicting a frame from a previously-decoded frame, by leveraging temporal correlations. In this paper, we propose an end-to-end learned system for compressing video frames. Instead of relying on pixel-space motion (as with optical flow), our system learns deep embeddings of frames and encodes their difference in latent space. At decoder-side, an attention mechanism is designed to attend to the latent space of frames to decide how different parts of the previous and current frame are combined to form the final predicted current frame. Spatially-varying channel allocation is achieved by using importance masks acting on the feature-channels. The model is trained to reduce the bitrate by minimizing a loss on importance maps and a loss on the probability output by a context model for arithmetic coding. In our experiments, we show that the proposed system achieves high compression rates and high objective visual quality as measured by MS-SSIM and PSNR. Furthermore, we provide ablation studies where we highlight the contribution of different components.

General information

Publication status: Published

MoE publication type: A4 Article in a conference publication

Organisations: Computing Sciences, Research group: Artificial Intelligence and Vision - AIV, Nokia Technologies

Contributors: Zou, N., Zhang, H., Cricri, F., Tavakoli, H. R., Lainema, J., Aksu, E., Hannuksela, M., Rahtu, E.

Number of pages: 5

Pages: 580-584

Publication date: 2020

Host publication information

Title of host publication: Proceedings - 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition

Workshops, CVPRW 2020

Publisher: IEEE

ISBN (Print): 978-1-7281-9361-8

ISBN (Electronic): 9781728193601

Publication series

Name: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops

ISSN (Print): 2160-7508

ISSN (Electronic): 2160-7516

ASJC Scopus subject areas: Computer Vision and Pattern Recognition, Electrical and Electronic Engineering

DOIs:

10.1109/CVPRW50498.2020.00079

Bibliographical note

EXT="Zhang, Honglei"

EXT="Cricri, Francesco"

EXT="Lainema, Jani"

JUFOID=70566

Source: Scopus

Source ID: 85090145918

Research output: Chapter in Book/Report/Conference proceeding › Conference contribution › Scientific › peer-review