33. International Conference in Central Europe
on
Computer Graphics, Visualization and Computer Vision 2025

in cooperation with the Eurographics Association, ACM and SIGGRAPH listed

www.WSCG.cz www.WSCG.eu

Primavera Hotel & Congress Center
Pilsen (Plzen), Czech Republic

26 - 29 May 2025

Conference Program

Conference Chairs

Vaclav Skala
University of West Bohemia, Plzen,
University of Defence, Brno
Czech Republic
www.VaclavSkala.eu

Nadia Magnenat-Thalmann,
Dr. Hon. Causa, U. of Ottawa, and U. of Hannover
University of Geneva/MIRALab, Geneva,
Switzerland

Keynote speaker

Gianmarco Cherchi
University of Cagliari, Cagliari,
Italy

Registration - Primavera Hotel – conference registration desk

- **Monday** 18:00 - 20:00 (recommended),
- **Tuesday-Wednesday** – ***only during breaks***

Conference DINNER on WEDNESDAY is in the city NOT in the hotel
Order ticket (donated) at the conference desk by Tuesday 10:00 - each 10 EUR.

Pilsen Self-guided Tour - http://wscg.zcu.cz/Plzen-Grand-Tour.pdf

Please, RELOAD the file to get updated schedule

Tuesday, May 27, 2025

8:00 – 8:30 Registration

9:00 – 9:10 – Opening session

9:10 - 10:40

- Schmidt, C., Overhoff, HM: PCA for Enhanced Cross-Dataset Generalizability in Breast Ultrasound Tumor Segmentation
  Paper code: [A71] Abstract: [A71]
- Madhavan, M., Nkhoma, P., Khoshkangini, R., Åberg, J., Jamali, M., Ljungqvist, M., Davidsson, P.: Object Detection Human Activity Recognition for Improved Patient Mobility and Caregiver Ergonomics
  Paper code: [A11] Abstract: [A11]
- Schwarz,A., Hernadi,L., Biessmann,F., Hildebrand,K.: The Influence of Faulty Labels in Data Sets on Human Pose Estimation
  Paper code: [A47] Abstract: [A47]
- Evain,A., Khemmar,R., Orzalesi,M., Ahmedali,S.: Improving 3D Monocular Object Detection with Dynamic Loss Function Adjustments
  Paper code: [C02] Abstract: [C02]

10:40 – 11:00 BREAK + Registration

11:00 – 11:50 Keynote Talk

- **Cherchi, Gianmarco: Robust Hexahedral Meshing with Grid-based approaches**
  Paper code: [A03] Abstract: [A03]

12:00 – 12:20 Common Photo

>> DO NOT MISS <<

12:20 - 14:00 Lunch BREAK

14:00 - 15:20

- Simon Leistikow, Vladimir Molchanov, Lars Linsen: Improving Comparability of Temporal Evolution in 2D Embeddings of Ensemble Data
  Paper code: [B47] Abstract: [B47]
- Komar, A., Barzegar Khalilsaraei, S., Augsdörfer, U.: 2D B-spline Curve Reconstruction using Convolutional Auto-Encoders and Distance Fields
  Paper code: [B29] Abstract: [B29]
- Raith,F., Scheuermann,G., Heine,C.: Simplifying Jacobi Sets Topology and Geometry by Selective Smoothing of Bivariate 2D Scalar Fields
  Paper code: [B71] Abstract: [B71]
- Halaj,M., Škorvánková,D., Madaras,M.: Metrically Accurate 3D Human Avatars from Silhouette Images
  Paper code: [B13] Abstract: [B13]

Break 15:20 – 15:50

15:50 – 17:00

- Koenen,J., Oehrl,S., Kuhlen,T., Gerrits,T.: Interactive Streaming of 3D Scenes to Mobile Devices using Dual-Layer Image Warping and Loop-based Depth Reconstruction
  Paper code: [A05] Abstract: [A05]
- Zöch, M.; Krispel, U.; Augsdörfer, U.: Learning 2D Triangular Meshes from Images with Transformers
  Paper code: [B23] Abstract: [B23]
- Cui, W., Wang, X., Fujisawa, M.: Extending Bonded DEM for Solid-Fluid Interaction: A Coupled BDEM-SPH Simulation Framework
  Paper code: [C37] Abstract: [C37]
- Garau, L., Cherchi, G.: A robust approach to detect intersections between triangles with different numerical representations
  Paper code: [A83] Abstract: [A83]
- Pokorny, P.: Visualization of Geographic Data in Blender
  Paper code: [A19] Abstract: [A19]

Wednesday, May 28, 2025

9:00 - 10:20

- Tolstykh,D., Slutskiy,D.: Balancing Bounding Box and Mask Annotations for Semi-Supervised Instance Segmentation
  Paper code: [B17] Abstract: [B17]
- Molchanov, V., Rave, H., Linsen, L.: Efficient Regularization-based Normalization for Interactive Multidimensional Data Analysis Without Scaling Artifacts
  Paper code: [B41] Abstract: [B41]
- Dhali.S., Dasgupta.B.: Occlusion SLAM: Improving Visual SLAM by Leveraging Occluded Points
  Paper code: [B83] Abstract: [B83]
- Paiola,P.H., Garcia,G.L., Manesco,J.R.R., Roder,M., Rodrigues,D., Papa,J.P.: Adapting LLMs for the Medical Domain in Portuguese: A Study on Fine-Tuning and Model Evaluation
  Paper code: [A13] Abstract: [A13]

10:20 – 10:40 BREAK

10:40 - 12:00

- Thißen, M., Tran, T. N. D., Esteve Ratsch, B., Schönbein, B. J., Trapp, U., Egner, B., Piat, R., Hergenröther, E.: Leveraging Large Language Models to Effectively Generate Visual Data for Canine Musculoskeletal Diagnoses
  Paper code: [A53] Abstract: [A53]
- Horváth, A. L.,Valasek, G., Bán, R.: Distance Fields in Monte Carlo Geometry Processing
  Paper code: [D19] Abstract: [D19]
- Saeedeh Barzegar Khalilsaraei, Ursula Augsdörfer: An Inception-based Variational Autoencoder for Curves Generation and Interpolation
  Paper code: [C83] Abstract: [C83]
- Justyna Niewiadomska-Kaplar: How to improve computer vision for color blindness: analytical and experimental approach to color vision of PIGOP model
  Paper code: [B73] Abstract: [B73]

12:00 – 14:00 LUNCH BREAK

14:00 – 16:00

- Kariko,Cs., Valasek,G.: Real-time Rendering of Algebraic Surfaces via Polynomial Fitting
  Paper code: [D07] Abstract: [D07]
- Krobath,D., Augsdoerfer,U.: Feature-Sensitive Mesh Simplification
  Paper code: [C67] Abstract: [C67]
- Jendoubi, F., Khemmar, R., Rossi, R., Haddad, M.: Assessing the Impact of Image Quality on Multi-Task Learning Models for 3D Object Detection and Drivable Area Segmentation
  Paper code: [C53] Abstract: [C53]
- Svirhunenko,Y., Tytarchuk,P., Holovko, Y., Tereschenko,Y.: HCTSketch: Hybrid CNN-Transformer Approach for Stroke Segmentation in Sketches
  Paper code: [C03] Abstract: [C03]
- Vad,V., Valasek,G.: Cubic Hermite Interpolations of Signed Distance Fields
  Paper code: [D23] Abstract: [D23]
- Gergely,G., Csaba, B.: Ocean Rendering with Fast Fourier Transform for Real-Time Applications
  Paper code: [C59] Abstract: [C59]

Closing session

Online presentations are on Wednesday – no physical session room

Conference Dinner – starts 19:00 (expected) – details at the registration desk.
Order ticket discounted on Tuesday morning
Place to be announced

Thursday, May 29, 2025

Online presentations only – no session in a physical room scheduled

(the order within a session might change due to technical problems)
Presentation: 15 mins. + 5 mins. discussion,
Platform: **ZOOM** – connection data were sent to the corresponding authors

9:00 - 10:00 - online presentations

- Rudolf,O., Hecker,R., Thißen,M, Sillekens,L., Penner,I., Seyfarth, S., Akelbein,J.-P., Hergenröther,E.: Implementation of visual people counting algorithms in embedded systems
  Paper code: [A59] Abstract: [A59]
- Ahmed Tabia .. Samia bouchafaa .. Fabien Bonardi: 6-DOF Pose Estimation For Event Cameras Using A Transformer-Based Approach
  Paper code: [B31] Abstract: [B31]
- Nikolov, I.: Preliminary Study of a Non-Direct Generative Image Anonymization Pipeline for Anomaly Detection
  Paper code: [B53] Abstract: [B53]

10:00 – 11:00 - online presentations

- Chang,T.V., Hartmann,K., Kuth,B., Seibt,S., von Rymon Lipinski,B.: EGO-VC – Evolutionary GPU-Optimization of Visual Correspondences for Image Alignment
  Paper code: [B02] Abstract: [B02]
- Raseta,M., Lesar,Z., Bohak,C.: Real-time volumetric explosion special effects for games
  Paper code: [D02] Abstract: [D02]
- Grinberg,Y., Harel,N., Goldberger,J., Lindenbaum,O.: Detect and Correct: A Selective Noise Correction Method for Learning with Noisy Labels
  Paper code: [C29] Abstract: [C29]
- Kordt,J.,Shekhar,S.,Lippert,C.: Efficient Semi-automatic Segmentation-Labeling of Any Volumetric Medical Image
  Paper code: [B19] Abstract: [B19]

11:00 – 12:00 BREAK

12:00 – 13:00 - online presentations

- Abdul Gafoor,M., Preda,M., Zaharia,T.: Refining Gaussian Splatting: A Volumetric Densification Approach
  Paper code: [A79] Abstract: [A79]
- Ljubic,J., Bohak,C.: Automated Conversion of 3D Mesh Car Models into LEGO Brick Sets Using Voxel-Based Optimization
  Paper code: [D03] Abstract: [D03]
- Hegazy,E., Gabr,M.: Enhancing Image Thresholding with Masi Entropy: An Empirical Approach to Parameter Selection
  Paper code: [B97] Abstract: [B97]

13:00 – 14:20 - online presentations

- Medina E., Loh L.: PeDesCar: A Large-Scale Dataset for Simulated Car-Pedestrian Collisions to Advance Safety Research
  Paper code: [C47] Abstract: [C47]
- Susanne Wulz, Ulrich Krispel: Detecting Out-Of-Distribution Labels in Image Datasets With Pre-trained Networks
  Paper code: [C41] Abstract: [C41]
- Tiaya Tedonchio,C., Rivest,L., Lemieux,P.-O.: Towards a grasping-oriented categorization of rigid industrial objects to support automatic grasping of 3D CAD models in virtual ergonomics
  Paper code: [C79] Abstract: [C79]
- Nuzhdin,A., Nagaev,A., Sautin,A., Kapitanov,A., Kvanchiani,K.: HaGRIDv2: 1M Images for Static and Dynamic Hand Gesture Recognition
  Paper code: [A29] Abstract: [A29]

14:00 – 15:40 - online presentations

- Karpowicz,F., Kepinski,W., Staszynski,B., Sarwas,G.: Examination of PCA Utilisation for Multilabel Classifier of Multispectral Images
  Paper code: [C97] Abstract: [C97]
- Weich, P., Lobachev, O.: PIN-a-Boo: Revealing Smartphone PINs via Segmentation and Hand Skeleton Tracking from Video Feeds
  Paper code: [D29] Abstract: [D29]
- Hegazy,E., Gabr,M.: A multi-level thresholding algorithm for threshold count and values identification based on dynamic programming
  Paper code: [B89] Abstract: [B89]
- Fatima Basim Jasim , Ali Retha Hasoon Khayeat: Object Detection and Recognition in Low Light Image Using Deep Learning Techniques
  Paper code: [A17] Abstract: [A17]
- Kaczmarek AL., Bierylo K., Gabryolek J., Macioch J.: Is Python OpenCV slower than its C++ version?
  Paper code: [A07] Abstract: [A07]
- Khapre,S., Semwal,S.: Mitigation of Social Media Platforms Impact on the Users
  Paper code: [B79] Abstract: [B79]
- Ananthakrishnan,A., Anush,B.,Dharanivendhan,V.,Ramanathan,M.,: CAD-RAG: A multi-modal retrieval augmented framework for user editable 3D CAD model generation
  Paper code: [D11] Abstract: [D11]
- Kvanchiani,K., Kraynov,R., Petrova,E., Surovcev,P., Nagaev,A., Kapitanov,A.: Training Strategies for Isolated Sign Language Recognition
  Paper code: [A37] Abstract: [A37]

WSCG 2025 – ABSTRACTS

Keynote

A03: Robust Hexahedral Meshing with Grid-based approaches

Cherchi, Gianmarco

Abstract:
Conforming hexahedral meshes are considered a prominent computational domain for simulation tasks due to their nice numerical properties. Despite their advantages, automatically decomposing a general 3D volume into a minimal number of hexahedral elements remains extremely challenging. One of the most effective strategies to address this challenge involves creating an adaptive Cartesian grid, which is then converted into a conforming hexahedral mesh. These methods are recognized for their robustness and are the only ones concretely used in the industry. The most advanced tools for this task are based on mesh dualization. This approach employs topological schemes to regularize the valence of internal vertices and edges, ensuring that the dualization process results in a pure hexahedral mesh. In this talk, we review the main aspects of grid-based approaches for generating hexahedral meshes, starting from the pioneering work of Marechal [Mar09] to the most recent advancements in the field, focusing on [Gao19], [Liv21] and [Pit21].

FULL papers

A05: Interactive Streaming of 3D Scenes to Mobile Devices using Dual-Layer Image Warping and Loop-based Depth Reconstruction

Koenen,J., Oehrl,S., Kuhlen,T., Gerrits,T.

Abstract:
While mobile devices have developed into hardware with advanced capabilities for rendering 3D graphics, they commonly lack the computational power to render large 3D scenes with complex lighting interactively. A prominent approach to tackle this is rendering required views on a remote server and streaming them to the mobile client. However, the rate at which servers can supply data is limited, e.g., by the available network speed, requiring image-based rendering techniques like image warping to compensate for the latency and allow a smooth user experience, especially in scenes where rapid user movement is essential. In this paper, we present a novel streaming approach designed to minimize artifacts during the warping process by including an additional visibility layer that keeps track of occluded surfaces while allowing access to 360Â° views. In addition, we propose a novel mesh generation technique based on the detection of loops to reliably create a mesh that encodes the depth information required for the image warping process. We demonstrate our approach in a number of complex scenes and compare it against existing works using two layers and one layer alone. The results indicate a significant reduction in computation time while achieving comparable or even better visual results when using our dual-layer approach.

A11: Object Detection Human Activity Recognition for Improved Patient Mobility and Caregiver Ergonomics

Madhavan, M., Nkhoma, P., Khoshkangini, R., Aberg, J., Jamali, M., Ljungqvist, M., Davidsson, P.

Abstract:
This study explores the use of machine learning to enhance patient mobility and caregiver ergonomics by optimizing the use of mobility aids. Traditional manual assessments can be subjective and inaccurate, so this research develops a data-driven model for object detection and human activity recognition. A computer vision dataset was created using video recordings of controlled caregiving scenarios. The study leverages advanced machine learning models, including YOLO for object detection, pose estimation, ResNet-18 for frame classification, Inception-v4 for feature extraction, and LSTM for sequence m deling. The findings provide valuable insights into integrating machine learning into mobility aids, improving both patient outcomes and caregiver well-being

A29: HaGRIDv2: 1M Images for Static and Dynamic Hand Gesture Recognition

Nuzhdin,A., Nagaev,A., Sautin,A., Kapitanov,A., Kvanchiani,K.

Abstract:
This paper proposes the second version of the widespread image-based Hand Gesture Recognition dataset HaGRID -- HaGRIDv2. We have added 15 new gestures to the existing 18, encompassing both conversational and control functions, including two-handed gestures. Building on the foundational concepts proposed by HaGRID"s authors, we implemented the dynamic gesture recognition algorithm and further enhanced it by adding three new groups of manipulation gestures. The „no gesture" class was significantly expanded and redefined with a new semantic focus to include a diverse range of natural hand movements, which led to a 16-fold reduction in false positives on HaGRIDv2. The HaGRIDv2 dataset outperforms the original HaGRID in pre-training models for gesture-related tasks. Besides, we achieved the best generalization ability among gesture and hand detection datasets. Additionally, the second version of the dataset provides a diverse range of hand samples, which is crucial for fine-tuning modern diffusion models. By fine-tuning on HaGRIDv2, these models achieve improved outcomes in generating anatomically correct hand gesture images. HaGRIDv2, pre-trained models, and a dynamic gesture recognition algorithm are publicly available.

A37: Training Strategies for Isolated Sign Language Recognition

Kvanchiani,K., Kraynov,R., Petrova,E., Surovcev,P., Nagaev,A., Kapitanov,A.

Abstract:
Accurate recognition and interpretation of sign language are crucial for enhancing communication accessibility for deaf and hard of hearing individuals. However, current approaches of Isolated Sign Language Recognition (ISLR) often face challenges such as low data quality and variability in gesturing speed. This paper introduces a comprehensive model training pipeline for ISLR designed to accommodate the distinctive characteristics and constraints of the Sign Language (SL) domain. The constructed pipeline incorporates carefully selected image and video augmentations to tackle the challenges of low data quality and varying sign speeds. Including an additional regression head combined with IoU-balanced classification loss enhances the model"s awareness of the gesture and simplifies capturing temporal information. Extensive experiments demonstrate that the developed training pipeline easily adapts to different datasets and architectures. Additionally, the ablation study shows that each proposed component expands the potential to consider ISLR task specifics. The presented strategies enhance recognition performance across various ISLR benchmarks and achieve state-of-the-art results on the WLASL and Slovo datasets.

A47: The Influence of Faulty Labels in Data Sets on Human Pose Estimation

Schwarz,A., Hernadi,L., Biessmann,F., Hildebrand,K.

Abstract:
In this study we provide empirical evidence demonstrating that the quality of training data impacts model performance in Human Pose Estimation (HPE). Inaccurate labels in widely used data sets, ranging from minor errors to severe mislabeling, can negatively influence learning and distort performance metrics. We perform an in-depth analysis of popular HPE data sets to show the extent and nature of label inaccuracies. Our findings suggest that accounting for the impact of faulty labels will facilitate the development of more robust and accurate HPE models for a variety of real-world applications. We show improved performance with cleansed data.

A53: Leveraging Large Language Models to Effectively Generate Visual Data for Canine Musculoskeletal Diagnoses

Thissen, M., Tran, T. N. D., Esteve Ratsch, B., Schoenbein, B. J., Trapp, U., Egner, B., Piat, R., Hergenroether, E.

Abstract:
More data generally improves AI model performance, but data collection can be challenging due to rarity or high costs. These challenges are evident in our effort to apply AI models to a novel approach for visually documenting the musculoskeletal condition of dogs, where palpated abnormalities are marked as colored strokes on a body map. Since these strokes correspond to distinct muscles or joints, they can be mapped to the textual domain in which large language models (LLMs) operate. LLMs have shown promise in medical applications, offering potential for synthetic training data generation. We investigate whether LLMs can effectively generate synthetic visual training data for canine musculoskeletal diagnoses. To achieve this, we developed a rendering process and a mapping that segments visual documentations into over 200 labeled regions representing muscle groups, muscles, or joints. Using techniques like guided decoding, chain-of-thought reasoning, and few-shot prompting, we generated 1,000 synthetic visual documentations for patellar luxation (kneecap dislocation) diagnosis, the diagnosis for which we have the most real-world data. Our analysis shows that the generated documentations accurately reflect diagnosis location and severity while remaining independent of the dog"s sex. However, they lack variation based on weight and age, highlighting some limitations. We further generated 1,000 visual documentations for various other diagnoses to create a binary classification dataset. A model trained solely on this synthetic data achieved an F1 score of 88% on 70 real-world documentations. These results highlight the potential of LLM-generated synthetic data to mitigate data scarcity, particularly for rare diseases.

B83: Occlusion SLAM: Improving Visual SLAM by Leveraging Occluded Points

Dhali.S., Dasgupta.B.

Abstract:
This paper presents a visual SLAM (simultaneous localization and mapping) system, called Occlusion SLAM, that uses occluded pointsâ€”map points that are sometimes visible and sometimes blocked by objects in the fore- groundâ€”and demonstrates a noticeable improvement in camera motion estimation. Occlusion SLAM is built upon ORB SLAM2, with an additional segmentation module based on YOLOv8 and the proposed occluded point detection module. The segmentation module creates a binary mask image of the predefined dynamic objects, which the occlusion detection module uses to identify occluded points, as they are intermittently revealed by the motion of an object. Occlusion SLAM was evaluated on highly dynamic datasets, such as the "walking" and "crowd" sequences from the TUM and Bonn datasets, respectively. Results indicate that the occluded point detection module, combined with segmentation, achieves significant improvements in absolute trajectory estimation with minimal computational requirements compared to ORB SLAM2. The method showed its full potential when at least one-fifth of the map points in a keyframe were identified as occluded points. The experiments also show competitive performance against other well known methods, such as DS SLAM, Dyna SLAM, RDS SLAM, Crowd SLAM and SG SLAM.

A59: Implementation of visual people counting algorithms in embedded systems

Rudolf,O., Hecker,R., Thissen,M, Sillekens,L., Penner,I., Seyfarth, S., Akelbein,J.-P., Hergenroether,E.

Abstract:
Optimising the efficiency of HVAC systems represents a significant opportunity to reduce energy consumption in buildings and mitigate greenhouse gas emissions. This research evaluates low-resolution computer Vision algorithms for occupancy detection on resource-constrained embedded systems. Our evaluation focuses specifically on the feasibility of deploying advanced AI object detection models on low-cost hardware platforms (under â‚¬10) with varying computational capabilities. We systematically compared 45 different pre-trained object detection models using the COCO dataset. Among the models evaluated, those with the YOLO backbone proved to be the most suitable for this task. Quantitative analysis showed that YOLOv5n achieved a favourable balance between accuracy (AP50 = 0.944; AP50-95 = 0.584), model size (2.6MBin RKNN format) and inference time. Performance tests on three embedded platforms - ESP32-CAM (microcontroller), Raspberry Pi Zero 2 W and Luckfox Pico Mini A (single-board computer) - revealed significant differences in inference speed, with hardware-accelerated solutions up to 10,000 times faster than software-only implementations. We have verified real-world applicability using our own ceiling-mounted wide-angle camera dataset. Future work will focus on developing a full Hardware prototype, optimising the training dataset with AI-generated synthetic data, and implementing sensor fusion with audio signals for a multimodal approach.

A71: PCA for Enhanced Cross-Dataset Generalizability in Breast Ultrasound Tumor Segmentation

Schmidt, C., Overhoff, HM

Abstract:
In medical image segmentation, limited external validity remains a critical obstacle when models are deployed across unseen datasets, an issue particularly pronounced in the ultrasound image domain. Existing solutions-such as domain adaptation and GAN-based style transfer-while promising, often fall short in the medical domain where datasets are typically small and diverse. This paper presents a novel application of principal component analysis (PCA) to address this limitation. PCA preprocessing reduces noise and emphasizes essential features by retaining approximately 90% of the dataset variance. We evaluate our approach across six diverse breast tumor ultrasound datasets comprising 4,794 B-mode images and corresponding expert tumor segmentation masks. For each dataset, a corresponding dimensionality reduced PCA-dataset is created and U-Net-based segmentation models are trained on each of the twelve datasets. Each model trained on an original dataset was inferenced on the remaining five out of-domain original datasets (baseline results), while each model trained on a PCA dataset was inferenced on five out-of-domain PCA datasets. Our experimental results indicate that using PCA reconstructed datasets, instead of original images, improves the model's recall and Dice scores, particularly for model-dataset pairs where baseline performance was lowest, achieving statistically significant gains in recall (0.57 Â± 0.07 vs. 0.70 Â± 0.05, p=0.0004) and Dice scores (0.50 Â± 0.06 vs. 0.58 Â± 0.06, p = 0.03). Our method reduced the decline in recall values due to external validation by 33%. These findings underscore the potential of PCA reconstruction as a safeguard to mitigate declines in segmentation performance, especially in challenging cases, with implications for enhancing external validity in real-world medical applications. Future studies are proposed to optimize PCA configurations for diverse imaging datasets and exploring integration with existing external validation methods.

A79: Refining Gaussian Splatting: A Volumetric Densification Approach

Abdul Gafoor,M., Preda,M., Zaharia,T.

Abstract:
Achieving high-quality novel view synthesis in 3D Gaussian Splatting (3DGS) often depends on effective point primitive management. The underlying Adaptive Density Control (ADC) process addresses this issue by automating densification and pruning. Yet, the vanilla

3DGS densification strategy shows key shortcomings. To address this issue, in this paper we introduce a novel density control method, which exploits the volumes of inertia associated to each Gaussian function to guide the refinement process. Furthermore, we study the effect of both traditional Structure from Motion (SfM) and Deep Image Matching (DIM) methods for point cloud initialization. Extensive experimental evaluations on the Mip-NeRF 360 dataset demonstrate that our approach surpasses 3DGS in reconstruction quality, delivering encouraging performance across diverse scenes.

A83: A robust approach to detect intersections between triangles with different numerical representations

Garau, L., Cherchi, G.

Abstract:
The detection and classification of intersections between triangles are crucial tasks in a wide range of applications within Computer Graphics and Geometry Processing, including mesh Arrangements, mesh Booleans, and generic mesh processing and fixing tasks. Existing methods are hard-coded and deeply integrated into specific algorithms, and significant efforts are usually required to integrate them into new pipelines or to extend them to different numerical representations. This paper presents a versatile and exhaustive algorithm to identify and classify intersections between triangles with either floating points, rational numbers, or implicit representations. The proposed tool is implemented as a C++ templated and header-only code that is generic and easy to integrate into further algorithms requiring the triangle-triangle intersection detection step. The developed tool has been tested and compared with a state-of-the-art approach, and it is shared with the Geometry Processing community with an Open Source license.

B02: EGO-VC â€" Evolutionary GPU-Optimization of Visual Correspondences for Image Alignment

Chang,T.V., Hartmann,K., Kuth,B., Seibt,S., von Rymon Lipinski,B.

Abstract:
Many computer vision applications, such as panoramic stitching, image morphing, and image registration depend on precise feature correspondences. While current machine-learning-based methods excel at detecting precise feature matches, they do not necessarily guarantee visual alignment. This work presents an evolutionary, mesh-based optimization framework that enhances alignment quality and improves visual coherence as a postprocessing step for any feature-matching algorithm: The approach employs the well-known enhanced correlation coefficient (ECC) as a visual error metric, efficiently computed in parallel using the rasterization capabilities of modern graphics hardware. Given any state-of-the-art feature matcher, the proposed method constructs a Delaunay feature point mesh and evolutionarily refines the ECC alignment of each generated triangle. This results in a more precise registration beyond the accuracy of initial feature matches. Extensive evaluation across multiple standard image datasets confirms the proposed method"s effectiveness, yielding ECC improvements up to 18.3% and ensuring high visual quality in downstream applications, particularly in challenging image areas with occlusions. For morphing-based applications, this leads to sharper, smoother transitions between image pairs, minimizing visual artifacts.

B13: Metrically Accurate 3D Human Avatars from Silhouette Images

Halaj,M., Skorvankova,D., Madaras,M.

Abstract:
In recent years, the demand for realistic human avatars has escalated across diverse industries, ranging from gaming and virtual or augmented reality to fashion and healthcare. Creating an accurate, lifelike virtual duplicate of a human subject requires a precise reconstruction of anthropometric measurements from the real body to the rendered body model. In this paper, we propose a novel pipeline for 3D human body model reconstruction from a set of two input images. Our method generates metrically accurate virtual human avatars, based on body measurements extracted from an input image. Our approach is

capable of parameterizing the generated mesh with body measurements and yields accurate results. To augment the metrically accurate body meshes, we introduce a pipeline for generating textured clothes to enhance user virtual experience. As a part of our work, we generated a synthetic dataset that serves as a foundation for further augmentation of the training process when extracting anthropometric measurements from an input photo. This dataset contains over one million files derived from 120,000 virtual avatars that can extend existing real datasets.

B17: Balancing Bounding Box and Mask Annotations for Semi-Supervised Instance Segmentation

Tolstykh,D., Slutskiy,D.

Abstract:
Instance segmentation models are crucial for precise object detection but often require expensive pixel-wise mask annotations. This paper studies the impact of combining bounding box and mask annotations in semi-supervised segmentation. We propose a method that leverages from both types of labeled data within a unified training framework. Through experiments on YOLO (convolution-based) and DETR (transformer-based) architectures, we demonstrate that balancing these annotation types significantly enhances performance while reducing labeling costs, particularly in terms of manual annotation time. Additionally, we evaluate few-shot and zero-shot scenarios, further highlighting the flexibility and efficiency of our method for budget-constrained segmentation tasks.

B19: Efficient Semi-automatic Segmentation-Labeling of Any Volumetric Medical Image

Kordt,J.,Shekhar,S.,Lippert,C.

Abstract:
Regions of interest are often labeled in volumetric medical images either for research purposes or for diagnosis and treatment planning. However, labeling such segments manually is time-consuming and requires medical expertise, which makes it expensive. We design a novel semi-automatic 3D workflow which allows efficient segmentation-labeling of volumetric images. To this end, for a given 3D image we first manually label a subset of its 2D slices using MedSAM (a foundational model for segmenting any 2D medical image) via bounding-box prompting. Subsequently, we interpolate user-provided prompts for the remaining slices to automatically generate labels for them. This way, users can process a complete volumetric image while working on only a subset of its slices. We evaluate our method on the diverse set of medical image datasets from the Medical Segmentation Decathlon challenge. Our approach significantly reduces the labeling effort, around 67%, while only marginally reducing the segmentation accuracy compared to applying MedSAM slice-by-slice. Breaking out of the time-consuming slice-by-slice workflow with only a minor reduction in accuracy is a significant step in streamlining the process of semi-automatic labeling. We will make the source code publicly available.

B23: Learning 2D Triangular Meshes from Images with Transformers

Zoech, M.; Krispel, U.; Augsdoerfer, U.

Abstract:
In this work, we present a method for learning 2D triangular meshes from images using the Transformer architecture. Creating polygonal representations of 2D surfaces from images has a wide range of applications, e.g. texture mapping and image segmentation. Current machine learning based methods either rely on deforming template meshes with fixed topology or require expensive post-processing to extract a planar polygonal mesh. In this paper, we demonstrate that deep learning can be utilized to directly generate a planar polygonal surface from arbitrary images without the need for additional input or constraints. Specifically, we show that the attention mechanism in transformer networks is highly effective in learning a unified representation of vertex positions and their neighborhood relationships. To showcase this, we propose a self-supervised training

pipeline that enables end-to-end learning of meshes directly from images. We apply our approach to various datasets of handwritten letters and digits, and show that the model is capable of learning meshes with varying genus.

B29: 2D B-spline Curve Reconstruction using Convolutional Auto-Encoders and Distance Fields

Komar, A., Barzegar Khalilsaraei, S., Augsdoerfer, U.

Abstract:
Curve reconstruction is a crucial task in various research domains, including Computer-Aided Design (CAD) and Reverse Engineering. Recovering a B-spline control polygon from a given set of points representing a curve remains an active area of study. We introduce a novel approach that leverages a distance field representation of the curve as input to two neural networks to reconstruct a closed B-spline. One network predicts distance fields to the control points, while the other estimates distances to the control polygon. Using these outputs, we determine the connectivity between control points, enabling the reconstruction of the B-spline control polygon. Our method is evaluated against state-of-the-art machine learning techniques and traditional optimization-based approaches.

B31: 6-DOF Pose Estimation For Event Cameras Using A Transformer-Based Approach

Ahmed Tabia, Samia Bouchafaa, Fabien Bonardi

Abstract:
Event cameras are novel sensors that provide significant advantages over traditional cameras, such as low latency, high dynamic range, and reduced motion blur. These properties make them particularly well-suited for 6-DOF pose estimation tasks in challenging environments. In this paper, we present a novel transformer-based approach for 6-DOF pose estimation using event camera data. Our method combines a pretrained ResNet50 backbone for feature extraction with a custom transformer encoder to model the spatial and temporal dependencies inherent in event data. We demonstrate the effectiveness of our approach on a dataset of real-world event camera images, where we achieve significant improvements in pose estimation accuracy compared to state-of-the-art methods. Additionally, our method exhibits robustness to varying lighting conditions, motion blur, and sensor noise, highlighting its potential for deployment in a wide range of applications, such as robotics, autonomous vehicles, and augmented reality. Our experimental results showcase the promising capabilities of transformer-based models in leveraging the unique properties of event cameras for accurate and efficient 6-DOF pose estimation.

B41: Efficient Regularization-based Normalization for Interactive Multidimensional Data Analysis Without Scaling Artifacts

Molchanov, V., Rave, H., Linsen, L.

Abstract:
Attribute values in multidimensional datasets often have different measurement units, making data normalization an essential preprocessing step for visualization algorithms such as multidimensional data projections. However, existing normalization techniques are often sensitive to noise, rely on specific data models, are computationally expensive, or have other limitations. The state-of-the-art method for computing optimal scalings of multidimensional data attributes is based on Lloyd relaxation in a linearly projected space. However, its high computational complexity hinders its applicability to datasets of moderate or large sizes. We overcome this limitation by efficiently regularizing the distribution of projected samples using integral images. Our method reduces scaling-induced artifacts, leading to more reliable multidimensional data analysis. Numerical experiments show that our approach outperforms or, at the very least, matches state-of-the-art methods in computation time, scalability, accuracy, and stability.

B47: Improving Comparability of Temporal Evolution in 2D Embeddings of Ensemble Data

Simon Leistikow, Vladimir Molchanov, Lars Linsen

Abstract:
Ensemble data refers to a set of simulations or measurements of the same phenomenon conducted under different settings such as varying control parameters of the simulation. When the data capture changes over time, the main analysis task is to compare the temporal evolution of the ensemble members. For a global analysis of the ensemble's temporal evolution, the data can be embedded in a 2D visual space, where each ensemble member is visualized by a 2D curve. For the embedding, the dissimilarities of the states of the ensemble members at different points in time shall be represented by the distances of the respective points in the 2D embedding. This goal is achieved by minimizing the stress functional of Multidimensional Scaling (MDS). However, direct visual comparison of the embedded curves can be complicated due to their non-trivial geometry. We propose an algorithm for the simplification of curve geometries in 2D MDS embeddings of ensemble data, where the target shape of a selected reference curve can be chosen by the user. We transform the original embedding in a controllable way consistent with the MDS stress functional. Thus, we improve the comparability of curves representing individual ensemble members, while minimizing the loss in accuracy of presented data dissimilarities. We perform several tests demonstrating the benefits of the proposed method, present our findings, and discuss numerical aspects of the algorithm.

B53: Preliminary Study of a Non-Direct Generative Image Anonymization Pipeline for Anomaly Detection

Nikolov, I.

Abstract:
With growing GDPR compliance demands for deep learning surveillance models, human anonymization is a key research area. Most studies use RGB images as input for generative models, which retain demographic features, compromising anonymization and consistency across frames. We present our initial study into a full-body anonymization pipeline for anomaly detection datasets, where the synthetic person generation never has access to the RGB pedestrian visuals. The proposed pipeline uses a combination of existing models for easier reproducibility. We use YoloV8 for object detection, ClipSeg and BiRefNet for segmentation, OpenPose for pose estimation, and an animation diffusion model to generate context-aware synthetic data. The final diffusion model processes only masks and skeletal pose images removing the problems with using sensitive data. We test on the Avenue dataset. We show that the proposed pipeline can consistently anonymize and change the demographic of detected pedestrians. We discuss the observed problems and the next steps in building a more robust second version.

B71: Simplifying Jacobi Sets Topology and Geometry by Selective Smoothing of Bivariate 2D Scalar Fields

Raith,F., Scheuermann,G., Heine,C.

Abstract:
The topological analysis of multivariate fields is vital when investigating the relationship between functions. Jacobi sets, the set of all points at which the gradients of the functions are linearly dependent, are an essential tool for such analyses, as they extend the notion of critical points from scalar fields to multivariate fields. However, the Jacobi sets can become very complex, in particular, due to numerical errors and noise. These problems occur in practice, such as in eddy detection on sea surfaces. Although several methods for simplifying Jacobi sets exist in the literature, they mainly reduce Jacobi sets visually without adjusting the function values, which is essential for further data processing. This paper introduces a novel algorithm that changes the values of functions in a 2D bivariate scalar field, resulting in simplified Jacobi sets. For this, we use a neighborhood graph to identify the Jacobi sets to simplify, visualize the complexity of the Jacobi set for real-world examples, and compare the results with prior work. The new approach preserves features better and simplifies the geometry of the Jacobi sets by reducing zigzag patterns.

B79: Mitigation of Social Media Platforms Impact on the Users

Khapre,S., Semwal,S.

Abstract:
Social Media Platforms offer numerous benefits and allow people to come together for various causes. Many communities, academia, government agencies, institutions, healthcare, entertainment, and businesses are on Social Media Platforms. They are intuitive and free for the users. It has become unimaginable to have life without social media. Their architecture and data handling are geared towards scalability, uninterrupted availability, and both personal and collaborative revenue generation. Primarily artificial Intelligence algorithms are employed on stored user data for optimization and feeds. This has the potential to impacts their usersâ€™ safety, privacy, and security even if meta data is used. A new decentralized data arrangement framework based on the Fractal-tree and L-Systems algorithm is proposed to mitigate some of the impacts of Social Media Platforms.

B89: A multi-level thresholding algorithm for threshold count and values identification based on dynamic programming

Hegazy,E., Gabr,M.

Abstract:
Multilevel image thresholding is a simple and efficient segmentation technique. Thresholding criteria such as Otsu and Kapur objective functions are extensively used in the literature. They are effective techniques but suffer from poor computational complexity. Thus, methods such as dynamic programming for exact optimization or metaheuristic algorithms for approximate optimization are applied to improve runtime. However, most of these algorithms take the count of thresholds as input. Hence, a novel input-less algorithm that can identify the count and values of thresholds simultaneously is proposed. The proposed method is then compared to state of the art methods to assess its efficiency and effectiveness.

B97: Enhancing Image Thresholding with Masi Entropy: An Empirical Approach to Parameter Selection

Hegazy,E., Gabr,M.

Abstract:
Image multilevel thresholding is used as a preprocessing step in computer vision or image processing applications. In recent years, various methods and criteria have been tested to improve thresholding performance and efficiency. State-of-the-art methods combine a metaheuristic optimization algorithm to optimize some objective function. Commonly used objective functions include Otsu criterion and entropy-based methods. Masi entropy is one of the objective functions that recently showed significant potential in image thresholding. It can deal with additivity and non expandability of information. However, in order for applying Masi entropy, there is a parameter r that needs to be tuned. Different works proposed different values and methods to determine the value of r. Nevertheless, there is a lack of validation and comparisons between these methods. This work is an experimental study validating a previously introduced method for selecting r. We compare the performance over two datasets including 700 various images while varying the number of thresholds from 1 to 15. Structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) metrics are used to evaluate the performance. The collective and individual performance improvements are elaborated. It is shown that the tested method achieves a significant improvement in segmentation quality over all tested numbers of thresholds. Higher numbers of thresholds showed greater improvement than smaller numbers. These results demonstrate the practical utility of the tested method in entropy-based image multilevel thresholding.

C02: Improving 3D Monocular Object Detection with Dynamic Loss Function Adjustments

Evain,A., Khemmar,R., Orzalesi,M., Ahmedali,S.

Abstract:
In 3D monocular object detection, optimizing the loss function is crucial for balancing multiple competing metrics, such as depth estimation, orientation, and object dimensions. Traditional approaches use a weighted sum of individual losses, allowing metric prioritization but risking training instability due to competition between terms. To address this, we first experimented with different loss function configurations to see how different loss interactions could emphasize specific metrics. These initial results demonstrated that abrupt changes in loss functions cause significant precision drops, therefore we decided to try dynamic loss functions adjustment, using transition functions to gradually shift metric emphasis over the training process. Among the tested transition functions, the Smoothstep function had the best balance across all metrics, followed by the Linear function, while the Smootherstep function provided strong initial performance but was eventually outperformed. Our results suggest that controlled, smooth transitions between different loss functions can enhance training stability and final detection accuracy, providing a way to improve 3D object detection models without overhauling their architecture.

C03: HCTSketch: Hybrid CNN-Transformer Approach for Stroke Segmentation in Sketches

Svirhunenko,Y., Tytarchuk,P., Holovko, Y., Tereschenko,Y.

Abstract:
Stroke segmentation is a crucial task in computer vision, particularly for analyzing sketches and handwritten drawings. Traditional methods, such as contour-based and clustering approaches, often struggle with noise, limited adaptability, and lack of deep structural understanding. Modern deep learning approaches, including U-Net and Mask R-CNN, focus primarily on object segmentation rather than analyzing individual strokes. In this work, we propose a hybrid approach combining Convolutional Neural Networks (CNN) and Vision Transformer (ViT) to effectively segment strokes while preserving both local features and global relationships. Our approach leverages data augmentation techniques to enhance generalization and improve segmentation performance. Experimental results demonstrate that our method achieves high accuracy, with a sketch classification accuracy of 98.81% and a stroke segmentation accuracy of 94.62%. Compared to existing solutions, our approach provides superior performance and robustness across different drawing styles.

C29: Detect and Correct: A Selective Noise Correction Method for Learning with Noisy Labels

Grinberg,Y., Harel,N., Goldberger,J., Lindenbaum,O.

Abstract:
Falsely annotated samples, also known as noisy labels, can significantly harm the performance of deep learning models. Two main approaches for learning with noisy labels are global noise estimation and data filtering. Global noise estimation approximates the noise across the entire dataset using a noise transition matrix, but it can unnecessarily adjust correct labels, leaving room for local improvements. Data filtering, on the other hand, discards potentially noisy samples but risks losing valuable data.
Our method identifies potentially noisy samples based on their loss distribution. We then apply a selection process to separate noisy and clean samples and learn a noise transition matrix to correct the loss for noisy samples while leaving the clean data unaffected, thereby improving the training process. Our approach ensures robust learning and enhanced model performance by preserving valuable information from noisy samples and refining the correction process. We applied our method to standard image datasets (MNIST, CIFAR-10, and CIFAR-100) and a biological scRNA-seq cell-type annotation dataset. We observed a significant improvement in model accuracy and robustness compared to traditional methods.

C37: Extending Bonded DEM for Solid-Fluid Interaction: A Coupled BDEM-SPH Simulation Framework

Cui, W., Wang, X., Fujisawa, M.

Abstract:
We present a novel fully Lagrangian simulation framework that couples Bonded Discrete Element Method (BDEM) with Smoothed Particle Hydrodynamics (SPH) to simulate the complex interactions between elastic solids and fluids. While particle-based methods have shown success in fluid-solid interaction simulations, most existing approaches focus on rigid or granular materials. Our framework extends this capability to elastic solids that can undergo deformations, fracture, and topological changes. The BDEM component represents solids as particles connected by elastic bonds that can stretch, bend, shear, twist and break, while SPH handles fluid dynamics with free surface flows. We introduce a parallel implementation strategy for bond computation that allows static bonds to be processed in parallel on GPU architectures. Our method naturally handles challenging phenomena such as fluid percolation through fractured solids and hydraulic fracturing without requiring complex interface tracking or remeshing operations. We demonstrate the framework"s effectiveness through various scenarios including elastic objects interacting with fluid flows, break-up of solids under hydraulic pressure, and objects with different elastic properties submerged in water. Our coupling particle-based framework naturally handles complex physical phenomena while maintaining computational efficiency, making it particularly suitable for computer graphics applications where visual plausibility and efficiency are prioritized over strict physical accuracy.

C41: Detecting Out-Of-Distribution Labels in Image Datasets with Pre-trained Networks

Susanne Wulz, Ulrich Krispel

Abstract:
Ensuring the correctness of annotations in training datasets is one way to increase the trustworthiness and reliability of Machine Learning. This study aims to detect semantic shifts in datasets using Feature-Based Out-Of-Distribution and outlier detection methods, assuming Out-Of-Distribution samples are far from In-Distribution data. The experiments began with distance-based methods, such as k-Nearest Neighbours and Mahalanobis, followed by feature pyramids and dimensionality reduction techniques to address high-dimensional challenges. The results showed that the k-Nearest Neighbours detector performed robustly, achieving 100\% AUROC when using ResNet50 on the Caltech-101 dataset, while the Mahalanobis detector showed unstable results with scores close to 50\%. Moreover, selecting the right backbone model and feature levels, particularly low-level features from ResNet50, improved performance achieving AUROC score of 96\% on the DelftBikes dataset for both k-Nearest Neighbours and Local Outlier Factor. The study highlights that k-Nearest Neighbours, Local Outlier Factor, alongside feature pyramids and dimensionality reduction constitute an effective setup for Out-of-Distribution detection, but optimal performance depends on tailored configurations across varying data conditions.

C47: PeDesCar: A Large-Scale Dataset for Simulated Car-Pedestrian Collisions to Advance Safety Research

Medina E., Loh L.

Abstract:
Car-pedestrian collisions are a daily occurrence worldwide, yet there is a notable absence of public datasets in this domain. Research in this area is crucial, as it directly impacts pedestrian safety and serves as a basis for validating autonomous driving systems. Although finite element simulations are used, they are computationally intensive and yield insufficient data for deep learning applications. In this work, we present the PeDesCar dataset for safe autonomous driving, which spans around 15 days of simulated time and encompasses over 1 million collision events, each constrained within a temporal window of up to 2 seconds per event. The dataset is generated using MuJoCo as a physics simulator, which has proven its effectiveness in sim2real robotics research. We use PeDesCar to train and assess state-of-the-art models in human motion prediction and validate the realism of

the simulation against realistic high-fidelity finite element simulations. Our results validate that PeDesCar is sufficient for preliminary car-pedestrian collision research. The visualization code, videos, and dataset are accessible on the project website.

C53: Assessing the Impact of Image Quality on Multi-Task Learning Models for 3D Object Detection and Drivable Area Segmentation

Jendoubi, F., Khemmar, R., Rossi, R., Haddad, M.

Abstract:
In autonomous driving, the quality of input images is crucial for the accuracy and reliability of perception systems, particularly in tasks like 3D object detection and drivable area segmentation. This study examines the impact of training Multi-Task Learning (M-TL) models exclusively on high-quality images for these applications. Using the KITTI dataset, we apply AI-based and traditional Image Quality Assessment (IQA) algorithms to filter and retain only high-quality images during training. Our experiments reveal that models trained on high-quality images achieve significantly better performance than those trained on the full dataset, including images of varying quality. These findings highlight the critical role of image quality in enhancing the accuracy and robustness of M-TL learning models for autonomous driving. Furthermore, this work emphasizes the importance of integrating image quality evaluation into the data-preprocessing pipeline to optimize model performance.

C67: Feature-Sensitive Mesh Simplification

Krobath,D., Augsdoerfer,U.

Abstract:
Feature recognition in meshes is a widely studied topic of computer science and is applied in order to identify and analyse properties of a 3D shape. It finds applications in numerous fields, like, e.g. medical imaging, robotics, cultural heritage, and CAD design. In this paper, we propose a new segmentation-based feature recognition approach for 3D meshes, where features are defined by the intersection of two neighboring segments. The detected features are used to derive very sparse polygonal representations of arbitrary input meshes, even merging features to achieve a high degree of simplification. We compare results for edge-based segmentation and a new normal-based segmentation and demonstrate how to derive good results even for pseudocurved surfaces and meshes affected by artifacts.

C79: Towards a grasping-oriented categorization of rigid industrial objects to support automatic grasping of 3D CAD models in virtual ergonomics

Tiaya Tedonchio,C., Rivest,L., Lemieux,P.-O.

Abstract:
The use of automatic analysis of 3D CAD models to identify a variety of plausible grasp locations on unknown objects with complex shapes and no affordances is a current concern in virtual ergonomics. In robotics, researchers approached this problem by grouping objects into categories related to geometric characteristics. Category-level grasping approaches have been little explored in virtual ergonomics. This paper introduces grasping-oriented categorization for rigid industrial objects to support the identification of a variety of plausible grasp locations in virtual ergonomics. First, we performed a rough analysis of objectsâ€™ convexity and define two high-level categories: global dominant shapes and local dominant shapes. Second, we propose a more detailed categorization of objects to sort them into four specific categories: 1- threadlike objects, 2- thin objects, 3- small objects, and 4- large and thick objects. Then, we explode those categories based on the presence or lack of holes or protrusions that are useful for grasping. This grasping-oriented categorization of objects is used to manually categorize a subset of 242 real rigid industrial objects. We obtained 27.3% threadlike objects, 43.8% thin objects, 2% small objects, 24.3% large and thick objects and 2.6% uncategorized objects. Our future work will pertain to develop approaches for automatically categorize unknown objects and approaches for

automatically identifying variety of plausible grasp locations on objects of given categories in virtual ergonomics.

C83: An Inception-based Variational Autoencoder for Curves Generation and Interpolation

Saeedeh Barzegar Khalilsaraei, Ursula Augsdoerfer

Abstract:
In this work, we introduce a novel approach for synthesizing new curves by leveraging Variational Autoencoder (VAE) latent space interpolation. Our method encodes existing ordered point sequences representing curves into a compact latent representation, enabling smooth and meaningful transitions between different curve shapes. By performing controlled interpolations in the learned latent space, we generate diverse, high-quality smooth curves that maintain structural coherence and geometric consistency. The proposed method is particularly useful for applications in shape design, procedural modeling, and data augmentation in geometric learning.

C97: Examination of PCA Utilisation for Multilabel Classifier of Multispectral Images

Karpowicz,F., Kepinski,W., Staszynski,B., Sarwas,G.

Abstract:
This paper investigates the utility of Principal Component Analysis (PCA) for multi-label classification of multispectral images using ResNet50 and DINOv2, acknowledging the high dimensionality of such data and the associated processing challenges. Multi-label classification, where each image may belong to multiple classes, adds further complexity to feature extraction. Our pipeline includes an optional PCA step that reduces the data to three dimensions before feeding it into a three-layer classifier. The findings demonstrate that the effectiveness of PCA for multi-label multispectral image classification depends strongly on the chosen deep learning architecture and training strategy, opening avenues for future research into self-supervised pre-training and alternative dimensionality reduction approaches.

D03: Automated Conversion of 3D Mesh Car Models into LEGO Brick Sets Using Voxel-Based Optimization

Ljubic,J., Bohak,C.

Abstract:
The transformation of 3D mesh car models into LEGO-compatible designs requires optimizing structural integrity, brick placement, and computational efficiency. This paper presents an algorithm specifically designed to convert polygonal car meshes into structurally sound LEGO representations. The approach involves voxelization, heuristic-based component mapping, and optimization techniques to ensure connectivity and maintain key design features. Using Blender for pre-processing and voxel-based methods for brick placement, the algorithm prioritizes larger LEGO elements to enhance stability and reduce part count. The proposed method is tested on various car models, demonstrating its ability to generate LEGO-compatible structures while preserving essential visual and functional details. The findings highlight potential applications in automated LEGO set design.

D07: Real-time Rendering of Algebraic Surfaces via Polynomial Fitting

Kariko,Cs., Valasek,G.

Abstract:
We investigate the problem of robust real-time algebraic surface rendering. We show that expressing the composition of the ray and the algebraic surface as a single univariate polynomial is not robust in practice, comparing results between monomial, Bernstein, Lagrange, and Chebyshev basis fits. We show that fitting multiple polynomials over subintervals, such as a unit length subdivision of the ray extent within the region of

interest, is a viable approach to overcome the robustness issues. In addition, we discuss a heuristic method for determining the split locations.

D11: CAD-RAG: A multi-modal retrieval augmented framework for user editable 3D CAD model generation

Ananthakrishnan,A., Anush,B.,Dharanivendhan,V.,Ramanathan,M.,

Abstract:
Computer-Aided Design (CAD) has revolutionized design and manufacturing by enabling precise, complex designs in collaborative environments. While similar CAD models with application-specific modifications are often needed, designs are typically created from scratch due to challenges in retrieving existing models or generating user-editable ones. Recent advancements in large foundational models like CLIP have enhanced cross-modal learning and retrieval, thereby improving search capabilities, but mostly trained on natural images which differs greatly from CAD domain. Additionally, parametric CAD modeling has been advanced through deep-generative learning approaches, where CAD is interpreted as a language task to generate user-editable designs. However, developing a multi-modal dataset tailored for 3D design tasks especially in engineering domain remains a challenge.

In this paper, the above ideas have been integrated, and a novel multi-modal pipeline is proposed for CAD command sequence generation using state-of-the-art Vision Language Models (VLMs). A unique multimodal CAD dataset, incorporating hand-drawn sketches, CAD command sequences, images, and basic text prompts, has been introduced. Subsequently, these modalities have been integrated through a Multi-modal Retrieval-Augmented Generation (MM-RAG) framework for user-editable CAD model retrieval and generation. The proposed RAG-based pipeline will aid designers by enabling iterative, user-editable CAD model generation based on simple user prompts. This approach aims to streamline the generation of CAD models, creating an advanced end-to-end pipeline designed to support designers.

D19: Distance Fields in Monte Carlo Geometry Processing

Horvath, A. L.,Valasek, G., Bain, R.

Abstract:
We propose to use discrete signed distance representations in Monte Carlo geometry processing simulations. In particular, we investigate algebraic and geometric generalizations of traditional signed distance fields. These are means to unify signed distance and closest point queries that are required by the walk-on-spheres algorithmic framework. We apply these to plane shapes enclosed by parametric polynomial boundaries. Our tests quantify the performance-accuracy trade-off compared to brute force closest point queries on test shapes.

D23: Cubic Hermite Interpolations of Signed Distance Fields

Vad,V., Valasek,G.

Abstract:
Signed Distance Functions are widely used in many fields of computer science, let alone computer graphics and modelling. In practice, a signed distance function values are computed on a regular grid, and later interpolation estimates the missing information. Recently, Hermite interpolation has been reported to be an effective method to overcome the weakness of highly used bilinear interpolation, while the extra derivative information must be computed/estimated. Despite the Hermite interpolation works well from a signal processing perspective, according to our knowledge there has been no study how these methods keep the Distance Field properties, such as (but not only) the Eikonal property.

D29: PIN-a-Boo: Revealing Smartphone PINs via Segmentation and Hand Skeleton Tracking from Video Feeds

Weich, P., Lobachev, O.

Abstract:
It is crucial to improve smartphone security, given the prevalence of sensitive information stored on them. This study presents an attack strategy that reveals smartphone PIN entries using computer vision and pattern recognition techniques. By leveraging modern segmentation and hand skeleton tracking, our method accurately identifies and analyzes finger movement patterns, even when partially obscured. We can reliably infer the entered PIN by combining these movement patterns with the smartphoneâ€™s position and the on-screen keypad layout. This approach significantly enhances shoulder-surfing attacks, requiring only a video recording of the entry process. Our attack requires much less specialized expertise, making it more accessible. We conclude by analyzing the methods potential impact and its implications for public safety.

SHORT PAPERS

A17: Object Detection and Recognition in Low Light Image Using Deep Learning Techniques

Fatima Basim Jasim , Ali Retha Hasoon Khayeat

Abstract:
Object detection in low-light conditions remains a critical challenge for applications in surveillance, autonomous navigation, security, etc. where poor image visibility results in poor detection accuracy especially in very low-light images. This paper presents an analysis of two popular object detection frameworks YOLOv9 and Faster R-CNN on low-light images after enhancing them using optimization techniques. We implemented a Proposed Method technique consisting of a set of image enhancement steps. Our results indicate that YOLOv9 and Faster R-CNN achieve superior performance with higher average recall, precision, and mAP. These results confirm the impact of integrating low-light enhancement methods on the detection accuracy in dealing with the complexities of low-light image environments. This paper contributes to the development of an image enhancement system for low-light conditions and the pursuit of better results in object detection in images.

B73: How to improve computer vision for color blindness: analytical and experimental approach to color vision of PIGOP model

Justyna Niewiadomska-Kaplar

Abstract:
The new theoretical approach on the structure of light radiations [Nie18a] that originated this research and the consequent theoretical description of the mechanisms of colour vision (PIGOP model), has produced the possibility of understanding the mechanisms of the formation of colour vision deficits in full agreement with the experimental results, which up to now have not found a coherent theoretical framework. In this article, considerations are made on how to improve computer vision for color blindness persons with the six most common types of colour vision deficiency, thus accommodating the needs of 8 percent of the world&quots male population.

D02: Real-time volumetric explosion special effects for games

Raseta,M., Lesar,Z., Bohak,C.

Abstract:
Video game special effects are usually created using sprites or simplistic triangle meshes with added textures, which is a good compromise between performance and quality. This paper presents an implementation of a real-time volumetric explosion effect that aims to

enhance realism in video games. The effect is rendered in a web environment using JavaScript and WebGPU. Our implementation uses ray marching for volume rendering, integrating a transfer function for color mapping, noise functions for adding natural variation, and bloom post-processing to enhance fire illumination. The contribution of this paper is to demonstrate the rendering possibilities and growing support of volumetric special effects for more interactive and immersive web applications and video games.

POSTER papers

A07: Poster: Is Python OpenCV slower than its C++ version?

Kaczmarek AL., Bierylo K., Gabryolek J., Macioch J.

Abstract:
The paper presents the comparison of the execution time of applications with regard to the programming language in which they were implemented. C++ and Python languages are considered. The selection of the programming language may have a significant impact on the overall performance of the application. The presented research applies to the OpenCV programming library which is a commonly used library in the field of image processing and computer vision. The paper presents experiments with Feature2D, Video I/O and highgui modules included in this library. In general it seems that programs written in C++ should be faster because it is a more low level language than Python. However surprisingly our research showed that in many cases using Python functions of OpenCV leads to the development of faster applications. The difference was particularly significant in case of the Feature2D module. In case of creating GUI it is negligible.

A13: Poster: Adapting LLMs for the Medical Domain in Portuguese: A Study on Fine-Tuning and Model Evaluation

Paiola,P.H., Garcia,G.L., Manesco,J.R.R., Roder,M., Rodrigues,D., Papa,J.P.

Abstract:
This study evaluates the performance of large language models (LLMs) as medical agents in Portuguese, aiming to develop a reliable and relevant virtual assistant for healthcare professionals. The HealthCareMagic-100k-en and MedQuAD datasets, translated from English using GPT-3.5, were used to fine-tune the ChatBode-7B model using the PEFT-QLoRA method. The InternLM2 model, with initial training on medical data, presented the best overall performance, with high precision and adequacy in metrics such as accuracy, completeness, and safety. However, DrBode models, derived from ChatBode, exhibited a phenomenon of catastrophic forgetting of acquired medical knowledge. Despite this, these models performed frequently or even better in grammaticality and coherence. A significant challenge was low inter-rater agreement, highlighting the need for more robust assessment protocols. This work paves the way for future research, such as evaluating multilingual models specific to the medical field, improving the quality of training data, and developing more consistent evaluation methodologies for the medical field.

A19: Poster: Visualization of Geographic Data in Blender

Pokorny, P.

Abstract:
This paper describes practical experience with the implementation and visualization of geographic data in Blender. Blender is a very popular open-source program primarily focused on 3D graphics, including a wide range of tools from model preparation and animation to final rendering. Given its popularity, it also implements support for a number of graphic formats from 2D and 3D graphics, as well as very useful add-ons that extend its functionality even further. One of these extensions is the implementation of geographic data that can be loaded from different sources. This contribution describes not only the actual ways of obtaining this data, but also how it is often modified to produce the best possible rendered visualizations.

C59: Poster: Ocean Rendering with Fast Fourier Transform for Real-Time Applications

Gergely,G., Csaba, B.

Abstract:
This paper presents simulation and rendering of ocean surfaces using modern graphics APIs. The simulation is based on the Jerry Tessendorf model, adapted for efficient parallel execution on the GPU. It employs analytical approximations of ocean wave spectra, generating a realistic ocean surface through Fourier transformations. We address the tiling artifacts caused by the simulation stage by combining multiple simulations with differing areas, creating a more dynamic surface. For rendering, we compare two approaches, one relies on patch tessellation and world partitioning to produce high-resolution geometry, and our other approach utilizes per-pixel displacement mapping to mimic fine details. We demonstrate a technique with a preprocessing cone mapping step that mitigates artifacts inherent to the latter algorithm.