WSCG 2023
31. International Conference on Computer Graphics. Visualization and Computer Vision

http://www.wscg.eu http://www.wscg.cz

Abstracts of accepted papers

Conference Chair
Prof. Vaclav Skala
www.VaclavSkala.eu
c/o Dept. of Computer Science and Engineering
Faculty of Applied Sciences, University of West Bohemia
Pilsen, Czech Republic

Venue
Primavera Hotel & Congress Centre
Pilsen, Czech Republic
May 15 – 18, 2023

- Monday 15 – Late afternoon & evening registration (exceptionally also during breaks)
- Tuesday – Wednesday - Thursday – Technical program
- Friday – departure (no scheduled presentations)

Pilsen, 12:00 - 2023-05-06

KEYNOTES

G43: Keynote: Illustrating Geometric Algebra and Differential Geometry in 5D Color Space

Benger,W.

Abstract:

Geometric Algebra (GA) is popular for its immediate geometric interpretations of algebraic objects and operations. It is based on Clifford Algebra on vector spaces and extends linear algebra of vectors by operations such as an invertible product, i.e. divisions by vectors. This formalism allows for a complete algebra on vectors same as for scalar or complex numbers. It is particularly suitable for rotations in arbitrary dimensions. In Euclidean 3D space quaternions are known to be numerically superior to rotation matrices and already widely used in computer graphics. However, their meaning beyond its numerical formalism often remains mysterious. GA allows for an intuitive interpretation in terms of planes of rotations and extends this concept to arbitrary dimensions by embedding vectors into a higher dimensional, but still intuitively graspable space of multi-vectors. However, out intuition of more than three spatial dimensions is deficient. The space of colors forms a vector space as well, though one of non-spatial nature, but spun by the primary colors red, green, blue. The GA formalism can be applied here as well, amalgamating surprisingly with the notion of vectors and co-vectors known from differential geometry: tangential vectors on a manifold correspond to additive colors red/green/blue, whereas co-vectors from the co-tangential space correspond to subtractive primary colors magenta, yellow, cyan. GA in turn considers vectors, bi-vectors and anti-vectors as part of its generalized multi-vector zoo of algebraic objects. In 3D space vectors, anti-vectors, bi-vectors and covectors are all three-dimensional objects that can be identified with each other, so their distinction is concealed. Confusions arise from notions such as normal vectors vs. axial vectors. Higher dimensional spaces exhibit the differences more clearly. Using colors instead of spatial dimensions we can expand our intuition by considering & transparency as an independent, four-dimensional property of a color vector. We can thereby explore 4D GA alternatively to spacetime in special/general relativity. However, even in 4D possibly confusing ambiguities remain between vectors, co-vectors, bi-vectors and bi-co-vectors: bi-vectors and bi-co-vectors - both six-dimensional objects - are visually equivalent.

They become unequivocal only in five or higher dimensions. Envisioning five-dimensional geometry is even more challenging to the human mind, but in color space we can add another property, texture to constitute a five-dimensional vector space. The properties of a bi-vector and a bi-co-vector becomes evident there: We can still study all possible combinations of colors/transparency/texture visually. This higher-dimensional yet intuitive approach demonstrates the need to distinguish among different kinds of vectors before identifying them in special situations, which also clarifies the meanings of algebraic objects in 3D Euclidean space and allows for better formulations of algorithms in 3D.

G47: Keynote: Raytracing Renaissance: An elegant framework for modeling light at Multiple Scales

Semwal,S,K.

Abstract:

Ray tracing remains of interest to Computer Graphics community with its elegant framing of how light interacts with objects, being able to easily support multiple light sources, and simple framework of merging synthetic and real cameras. Recent trends to provide implementations at the chip-level means raytracing’s constant quest of realism would propel its usage in real-time applications. AR/VR, Animations, 3DGames Industry, 3D-large scale simulations, and future social computing platforms are just a few examples of possible major impact. Raytracing is also appealing to HCI community because raytracing extends well along the 3D-space and time, seamlessly blending both synthetic and real cameras at multiple scales to support storytelling. This presentation will include a few milestones from my work such as the Slicing Extent technique and Directed Safe Zones. Our recent applications of applying machine learning techniques creating novel synthetic views, which could also provide a future doorway to handle dynamic scenes with more compute power as needed, will also be presented. It is once again renaissance for ray tracing which for last 50+ years has remained the most elegant technique for modeling light phenomena in virtual worlds at whatever scale compute power could support.

FULL Papers

D47: Automatic Individual Identification of Patterned Solitary Species Based on Unlabeled Video Data

Suessle,V., Arandjelovic,M., Kalan,A., Agbor,A., Boesch,C., Brazzola,G., Deschner,T., Dieguez,P., Granjon,A., Kuehl,H., Landsmann,A., Lapuente,J., Maldonado,N., Meier,A., Rockaiova,Z., Wessling,E., Wittig,R., Downs,C., Weinmann,A., Hergenroether,E.,

Abstract:

The manual processing and analysis of videos from camera traps is time-consuming and includes several steps, ranging from the filtering of falsely triggered footage to identifying and re-identifying individuals. In this study, we developed a pipeline to automatically analyze videos from camera traps to identify individuals without requiring manual interaction. This pipeline applies to animal species with uniquely identifiable fur patterns and solitary behavior, such as leopards (Panthera pardus). We assumed that the same individual was seen throughout one triggered video sequence. With this assumption, multiple images could be assigned to an individual for the initial database filling without pre-labeling. The pipeline was based on well-established components from computer vision and deep learning, particularly convolutional neural networks (CNNs) and scale-invariant feature transform (SIFT) features. We augmented this basis by implementing additional components to substitute otherwise required human interactions. Based on the similarity between frames from the video material, clusters were formed that represented individuals bypassing the open set problem of the unknown total population. The pipeline was tested on a dataset of leopard videos collected by the Pan African Programme: The Cultured Chimpanzee (PanAf) and achieved a success rate of over 83% for correct matches between previously unknown individuals. The proposed pipeline can become a valuable tool for future conservation projects based on camera trap data, reducing the work of manual analysis for individual identification, when labeled data is unavailable.

D53: ALIVE: Adaptive-Chromaticity for Interactive Low-light Image and Video Enhancement

Shekhar,S., Reimann,M., Wattasseril,J., Semmo,A., Döllner,J., Trapp,M.

Abstract:

Image acquisition in low-light conditions suffers from poor quality and significant degradation in visual aesthetics. This affects the visual perception of the acquired image and the performance of computer vision and image processing algorithms applied after acquisition. Especially for videos, the additional temporal domain makes it more challenging, wherein quality is required to be preserved in a temporally coherent manner. We present a simple yet effective approach for low-light image and video enhancement. To this end, we introduce Adaptive Chromaticity, which refers to an adaptive computation of image chromaticity. The above adaptivity avoids the costly step of low-light image decomposition into illumination and reflectance, employed by many existing techniques. Subsequently, we achieve interactive performance, even for high resolution images. Moreover, all stages in our method consists of only point-based operations and high-pass or low-pass filtering, thereby ensuring that the amount of temporal incoherence is negligible when applied on a per-frame basis for videos. Our results on standard low-light image datasets show the efficacy of our method and its qualitative and quantitative superiority over several state-of-the-art approaches. We perform a user study to demonstrate the preference for our method in comparison to state-of-the-art approaches for videos captured in the wild.

D59: LS-3DLane regression network using Parallelism loss function

Hassoubah,M., Sistu,G.

Abstract:

Accurate Lane position prediction is crucial in autonomous driving for safe vehicle maneuvering. Monocular cameras, aided by AI advancements, have proven to be effective in this task. However, 2D image space predictions overlook lane height, causing poor results in uphill or downhill scenarios. To address this issue, we present LS-3DLane network, inspired by the Lift-Splat-Shoot architecture, which predicts lane position in 3D space using a data-driven approach. The network also employs the Parallelism loss, using prior knowledge of lane geometry, to improve performance. Our results show that LS-3DLane outperforms previous approaches like Gen-LaneNet and 3D-LaneNet, with F-score improvements reaching 5.5\% and 10\% in certain cases. LS-3DLane performs similarly in X/Z error metrics.

D67: Visual Exploration of Repetitive Patterns on Ancient Peruvian Pottery

Lengauer,S.,Shao,L.,Mayerhofer,M.,Preiner,R.,Karl,S.,Trinkl,E.,Sipiran,I.,Bustos,B.,Schreck,T.

Abstract:

The analysis and understanding of artefact properties and their relationships is a key goal in archaeological analysis of cultural heritage objects. There are many aspects of concern, including shape properties of the objects, but also appearance properties stemming from paintings and ornamentations on the object surfaces. To date, these are considered by experts mostly holistically and on a per-object basis. We present an approach for the interactive visual exploration and correlation of shape- and ornament-based properties of a large collection of ancient vessels. Our approach allows to group objects by properties, and to relate them in side-by-side and bipartite graph displays. To this end, we define an encompassing set of feature descriptors, which are leveraged to cluster the objects by properties, selected by the user. Case studies show that a comparative overview of all objects effectively supports the discovery of interesting co-occurrences of shape and ornament properties. This way, our tool opens new possibilities for the domain analysis of cultural heritage object collections by data-driven visual exploration.

D73: Investigation on Encoder-Decoder Networks for Segmentation of Very Degraded X-Ray CT Tomograms

Dulau,I., Beurton-Aimar,M., Hwu,Y., Recur,B.

Abstract:
Field of View Nano-CT X-Ray synchrotron imaging is used for acquiring brain neuronal features from Golgi-stained bio-samples. It theoretically requires a large number of acquired radiographs for compensating reconstruction noise reinforced by the brain features sparsity. However reducing the number of radiographs is essential in routine applications but it results to degraded tomograms. In such a case, traditional segmentation methods are no longer able to distinguish neuronal structures from surrounding noise. We investigate several existing deep-learning networks and we define new ones to segment brain features from very degraded tomograms. We demonstrate the superiority of the proposed networks compared to existing ones.

D79: Self-Checkout Product Class Verification using Center Loss approach

Ciapas,B., Treigys,P.

Abstract:
The traditional image classifiers are not capable to verify if samples belong to specified classes due to several reasons: classifiers do not provide boundaries between in-class and out-of-class samples; although classifiers provide separation boundaries between known classes, classifiers" latent features tend to have high intra-class variance; classifiers often predict high probabilities for out-of-distribution samples; training classifiers on unbalanced data results in bias towards over-represented classes. The nature of the class verification problem requires a different loss function than the ubiquitous cross entropy loss in traditional classifiers: input to a class verification function includes a suggested class in addition to an image. As opposed to outlier detection, space is transformed to be not only separable, but discriminative between in-class and out-of-class inputs. In this paper, class verification based on a euclidean distance from the class centre is proposed and implemented. Class centres are learnt by training on a centre loss function. The method"s effectiveness is shown on a self-checkout image dataset of 194 food retail products. The results show that a two-fold loss function is not only useful to verify class, but does not degrade classification performance - thus, the same neural network is usable both for classification and verification.

D97: Why Existing Multimodal Crowd Counting Datasets Can Lead to Unfulfilled Expectations in Real-World Applications

Thissen,M., Hergenroether,E.

Abstract:
More information leads to better decisions and predictions, right? Confirming this hypothesis, several studies concluded that the simultaneous use of optical and thermal images leads to better predictions in crowd counting. However, the way multimodal models extract enriched features from both modalities is not yet fully understood. Since the use of multimodal data usually increases the complexity, inference time, and memory requirements of the models, it is relevant to examine the differences and advantages of multimodal compared to monomodal models. In this work, all available multimodal datasets for crowd counting are used to investigate the differences between monomodal and multimodal models. To do so, we designed a monomodal architecture that considers the current state of research on monomodal crowd counting. In addition, several multimodal architectures have been developed using different multimodal learning strategies. The key components of the monomodal architecture are also used in the multimodal architectures to be able to answer whether multimodal models perform better in crowd counting in general. Surprisingly, no general answer to this question can be derived from the existing datasets. We found that the existing datasets hold a bias toward thermal images. This was determined by analyzing the relationship between the brightness of optical images and crowd count as well as examining the annotations made for each dataset. Since answering this question is important for future real-world applications of crowd counting, this paper establishes criteria for a potential dataset suitable for answering whether multimodal models perform better in crowd counting in general.

E05: Coordinate-Unet 3D for segmentation of lung parenchyma

Van Linh,L., Olivier,S.

Abstract:
Lung segmentation is an initial step to provide accurate lung parenchyma in many studies on lung diseases based on analyzing the Computed Tomography (CT) scan, especially in Non-Small Cell Lung Cancer (NSCLC) detection.
In this work, Coordinate-UNet 3D, a model inspired by UNet, is proposed to improve the accuracy of lung segmentation in the CT scan. Like UNet, the proposed model consists of a contracting/encoder path to extract the high-level information and an expansive/decoder path to recover the features to provide the segmentation. However, we have considered modifying the structure inside each level of the model and using the Coordinate Convolutional layer as the final layer to provide the segmentation. This network was
trained end-to-end from a small set of CT scans of NSCLC patients. The experimental results show the proposed network can provide a highly accurate segmentation for the validation set with a Dice Coefficient index of 0.991, an F1 score of 0.976, and a Jaccard index (IOU) of 0.9535.

E07: Sex Classification of Face Images using Embedded Prototype Subspace Classifiers

Hast,A.

Abstract:
In recent academic literature Sex and Gender have both become synonyms, even though distinct definitions do exist. This give rise to the question, which of those two are actually face image classifiers identifying? It will be argued and explained why CNN based classifiers will generally identify gender, while feeding face recognition feature vectors into a neural network, will tend to verify sex rather than gender. It is shown for the first time how state of the art Sex Classification can be performed using Embedded Prototype Subspace Classifiers (EPSC) and also how the projection depth can be learned efficiently. The automatic Gender classification, which is produced by the emph{InsightFace} project, is used as a baseline and compared to the results given by the EPSC, which takes the feature vectors produced by emph{InsightFace} as input. It turns out that the depth of projection needed is much larger for these face feature vectors than for an example classifying on MNIST or similar. Therefore, one important contribution is a simple method to determine the optimal depth for any kind of data. Furthermore, it is shown how the weights in the final layer can be set in order to make the choice of depth stable and independent of the kind of learning data. The resulting EPSC is extremely light weight and yet very accurate, reaching over 98% accuracy for several datasets.

E17: Perceptions of Colour Pickers in Virtual Reality Art-Making

Alex, M., WÃ¼nsche, B., Lottridge, D.

Abstract:
Virtual reality art is reshaping digital art experiences, especially with the recent release of multiplayer 3D art applications, but may elicit different first impressions across different age groups which can impact their uptake. In particular, popular colour pickers based on HSV colour spaces may appeal differently to younger and older adults. We investigate first impressions of colour selection when shown with a discrete picker or a continuous HSV picker via an online survey with 63 adults and 24 older adults. We found that the discrete picker was seen as having more positive hedonic qualities overall; there were no differences between perceptions of adults and older adults. We discuss the implications of our findings for colour selection tools in virtual reality art-making.

E19: JengASL: A Gamified Approach to Sign Language Learning in VR

Shaw, A., Wunsche, B., Mariono, K., Ranveer, A., Xiao, M., Hajika, R., Liu, Y.

Abstract:
Learning sign language has many advantages ranging from being able to communicate with millions of hearing impaired people, to improving cognitive function and communication skills. Sign language is recognised as an official language in 74 countries, including Germany, Japan, and the UK. Despite that only a small percentage of people attempt to learn sign language.
In this research we investigate how virtual reality and gamification can be used to make learning sign language more enjoyable and motivating. We present JengASL, a gamified approach using 3D hand models, gesture recognition, and interactive gameplay in Virtual Reality to teach American Sign Language. We evaluate this system with a pilot study using eight participants and found that while it is less effective for sign memorisation than traditional 2D image-based learning methods, learning is more, but not significantly more, enjoyable and motivating.

E23: StarSRGAN: Improving Real-World Blind Super-Resolution

Vo,K.D, Bui,L.T.

Abstract:
The aim of blind super-resolution (SR) in computer vision is to improve the resolution of an image without prior knowledge of the degradation process that caused the image to be low-resolution. The State of the Art (SoA) model Real-ESRGAN has advanced perceptual loss and produced visually compelling outcomes using more complex degradation models to simulate real-world degradations. However, there is still room to improve the super-resolved quality of Real-ESRGAN by implementing recent research. This research paper introduces StarSRGAN, a novel GAN model designed for blind super-resolution (SR) tasks that utilizes various architectures. Our model provides SoA performance on the MANIQA and AHIQ measures, as demonstrated by experimental comparisons with past work. In addition, as a compact version, StarSRGAN Lite provides approximately 15 times faster reconstruction speed with acceptable quality, thereby facilitating the development of a real-time SR experience for future research.

E31: Modeling and Rendering with eXpressive B-Spline Curves

Seah,H.S., Tandianus,B., Sui,Y., Wu,Z., Zhang,Z.

Abstract:
eXpressive B-Spline Curve (XBSC) is a resolution-independent and computationally efficient technique for vector-based stroke modeling and rendering with the flexibility in defining and adjusting the shape and other parameters of the stroke. It generalizes the existing Disk B-Spline Curve (DBSC) geometric representation, which itself is a generalization of the Disk Bezier curve. XBSC allows flexible shape and color manipulation and rendering of strokes with asymmetrical shape control and rich color management. These properties make XBSC suitable for modeling freeform stroke shapes and animation, specifically in squash and stretch, a common technique to bestow elasticity and flexibility in shape changes. During the squash and stretch animation computation, we constrain the shape of the XBSC stroke to conserve its area. To achieve this, we apply the simulated annealing algorithm to iteratively adjust the XBSC while maintaining its area. We show several drawings, rendering and deformation examples to demonstrate the robustness of XBSC.

E41: Reconstruction from Multi-view Sketches: an Inverse Rendering Approach

Colom,J., Saito,H.

Abstract:
Reconstruction from real images has evolved very differently from reconstruction from sketches. Even though both present similarities, the latter aims to surpass the subjectivity that drawings present, increasing the task's uncertainty and complexity. In this work, we draw inspiration from reconstruction over real multi-view images and adapt existing methods in this area to work over sketches. We leverage inverse rendering as a refinement process for 3D colored meshes while proposing modifications for the domain of drawings. Compared to previous reconstructions from sketches, our proposal recovers not only shape but color, offering an optimization system that does not require previous training. Through the results, we evaluate how different quality factors in sketches affect the reconstruction quality and report how our proposal adapts to them compared to directly applying existing inverse rendering systems for real images.

E43: Training Image Synthesis for Shelf Item Detection reflecting Alignments of Items in Real Image Dataset

Tomokazu,K., Ryosuke,S., Soma,S.

Abstract:
We propose a novel cut-and-paste approach to synthesize a training dataset for shelf item detection, reflecting the alignments of items in the real image dataset. The conventional cut-and-paste approach synthesizes large numbers of training images by pasting foregrounds on background images and is effective for training object detection. However, the previous method pastes foregrounds on random positions of the background, so the alignment of items on shelves is not reflected, and unrealistic images are generated. Generating realistic images that reflect actual positional relationships between items is necessary for efficient learning of item detection. The proposed method determines the pasting positions for the foreground images by referring to the alignment of the items in the real image dataset, so it can generate more realistic images that reflect the alignment of the real-world items. Since our method can synthesize more realistic images, the trained models can perform better.

E47: SAIL: Semantic Analysis of Information in Light Fields: Results from Synthetic and Real-World Data

Kremer,R., Herfet,T.

Abstract:
Computational photography has revolutionized the way we capture and interpret images. Light fields, in particular, offer a rich representation of a scene"s geometry and appearance by encoding both spatial and angular information. In this paper, we present a novel approach to light field analysis that focuses on semantics. In contrast to the uniform distribution of samples in two-dimensional images, the distribution of samples in light fields varies for different scene regions. Some points are sampled from multiple directions, while others may only be captured by a small portion of the light field array. Our approach provides insights into this non-uniform distribution and helps guide further processing steps to fully leverage the available information content.

E53: Texture Spectral Similarity Criteria Comparison

Havlicek, M. Haindl, M.

Abstract:
Criteria capable of texture spectral similarity evaluation are presented and compared. From the fifteen evaluated criteria, only four criteria guarantee zero or minimal spectral ranking errors. Such criteria can support texture modeling algorithms by comparing the modeled texture with corresponding synthetic simulations. Another possible application is the development of texture retrieval, classification, or texture acquisition system. These criteria thoroughly test monotonicity and mutual correlation on specifically designed extensive monotonously degrading experiments.

E59: Designing a Lightweight Edge-Guided Convolutional Neural Network for Segmenting Mirrors and Reflective Surfaces

Gonzales,M.E.M., Uy,L.C., Ilao,J.P.

Abstract:
The detection of mirrors is a challenging task due to their lack of a distinctive appearance and the visual similarity of reflections with their surroundings. While existing systems have achieved some success in mirror segmentation, the design of lightweight models remains unexplored, and datasets are mostly limited to clear mirrors in indoor scenes. In this paper, we propose a new dataset consisting of 454 images of outdoor mirrors and reflective surfaces. We also present a lightweight edge-guided convolutional neural network based on PMDNet. Our model uses EfficientNetV2-Medium as its backbone and employs parallel convolutional layers and a lightweight convolutional block attention module to capture both low-level and high-level features for edge extraction. It registered maximum F-measure scores of 0.8483, 0.8117, and 0.8388 on the Mirror Segmentation Dataset (MSD), Progressive Mirror Detection (PMD) dataset, and our proposed dataset, respectively. Applying filter pruning via geometric median resulted in maximum F-measure scores of 0.8498, 0.7902, and 0.8456, respectively, performing competitively with the state-of-the-art PMDNet but with 78.20x fewer floating-point operations per second and 238.16x fewer parameters. The code and dataset are available at https://github.com/memgonzales/mirror-segmentation.

E67: Bias mitigation techniques in Image Classification: Fair Machine Learning in Human Heritage Collections

Ortiz Pablo,D., Badri,S., Noren,E., Notzli,C.

Abstract:
A major problem with using automated classification systems is that if they are not engineered correctly and with fairness considerations, they could be detrimental to certain populations. Furthermore, while engineers have developed cutting-edge technologies for image classification, there is still a gap in the application of these models in human heritage collections, where data sets usually consist of low-quality pictures of people with diverse ethnicity, gender, and age. In this work, we evaluate three bias mitigation techniques using two state-of-the-art neural networks, Xception and EfficientNet, for gender classification. Moreover, we explore the use of transfer learning using a fair data set to overcome the training data scarcity. We evaluated the effectiveness of the bias mitigation pipeline on a cultural heritage collection of photographs from the 19th and 20th centuries, and we used the FairFace data set for the transfer learning experiments. After the evaluation, we found that transfer learning is a good technique that allows better performance when working with a small data set. Moreover, the fairest classifier was found to be accomplished using transfer learning, threshold change, re-weighting and image augmentation as bias mitigation methods.

E71: A Resource Allocation Algorithm for a History-Aware Frame Graph

Sandu,R., Shcherbakov,A.

Abstract:
We consider the problem of memory consumption by a real-time GPU-accelerated graphical application. A history of a resource is defined for a particular frame to be the final contents of such a resource at the end of the previous frame. When organizing a graphical application using a frame rendering graph approach, it makes sense to implement automatic serving of resource history read requests of nodes. In absence of history resource requests, allocating resources for a fixed frame graph is the classic problem of dynamic storage allocation (DSA). In this paper, we formulate a generalization of DSA that enables memory reuse for resources with history requests and provide a practical approximate algorithm for solving it.

E79: Monte Carlo Based Real-Time Shape Analysis in Volumes

Gurijala,K.Ch., Wang,L., Kaufman,A.

Abstract:
We introduce a Monte Carlo based real-time diffusion process for shape-based analysis in volumetric data. The diffusion process is carried out by using tiny massless particles termed shapetons, which are used to capture the shape information. Initially, these shapetons are randomly distributed inside the voxels of the volume data. The shapetons are then diffused in a Monte Carlo fashion to obtain the shape information. The direction of propagation for the shapetons is monitored by the Volume Gradient Operator (VGO). This operator is known for successfully capturing the shape information and thus the shape information is well captured by the shapeton diffusion method. All the shapetons are diffused simultaneously and all the results can be monitored in real-time. We demonstrate several important applications of our approach including colon cancer detection and design of shape-based transfer functions. We also present supporting results for the applications and show that this method works well for volumes. We show that our approach can robustly extract shape-based features and thus forms the basis for improved classification and exploration of features based on shape.

E83: Probabilistic Pose Estimation with Synthetic-Real Domain Adaptation

Del-Tejo-Catala,O., Perez,J., Guardiola,J.L., Perez,A.J., Perez-Cortes,J.C.

Abstract:
Real samples are costly to acquire in many real-world problems. Thus, employing synthetic samples is usually the primary solution to train models that require large amounts of data. However, the difference between synthetically generated and real images, called domain gap, is the most significant hindrance to this solution, as it affects the model’s generalization capacity. Domain adaptation techniques are crucial to train models using synthetic samples. Thus, this article explores different domain adaptation techniques to perform pose estimation from a probabilistic multiview perspective. Probabilistic multiview pose estimation solves the problem of object symmetries, where a single view of an object might not be able to determine the 6D pose of an object, and it must consider its prediction as a distribution of possible candidates. GANs are currently state-of-the-art in domain adaptation. In particular, this paper explores CUT and CycleGAN, which have unique training losses that address the problem of domain adaptation from different perspectives. The datasets explored are a cylinder and a sphere extracted from a Kaggle challenge with perspective-wise symmetries, although they holistically have unique 6D poses. CUT outperforms CycleGAN in feature adaptation, although it is less robust than CycleGAN in keeping keypoints intact after translation, leading to pose prediction errors for some objects. Moreover, this paper found that training the models using synthetic-to-real images and evaluating them with real images improves the modelâ’s accuracy for datasets without complex features. This approach is more suitable for industrial applications to reduce inference overhead.

E89: Real-time Light Estimation and Neural Soft Shadows for AR Indoor Scenarios

Sommer, A., Schwanecke, U., Schoemer, E.

Abstract:
We present a pipeline for realistic embedding of virtual objects into footage of indoor scenes with focus on real-time AR applications. Our pipeline consists of two main components: A light estimator and a neural soft shadow texture generator. Our light estimation is based on deep neural nets and determines the main light direction, light color, ambient color and an opacity parameter for the shadow texture. Our neural soft shadow method encodes object-based realistic soft shadows as light direction dependent textures in a small MLP. We show that our pipeline can be used to integrate objects into AR scenes in a new level of realism in real-time. Our models are small enough to run on current mobile devices. We achieve runtimes of 9ms for light estimation and 5ms for neural shadows on an iPhone 11 Pro.

E97: A Framework for Art-directed Augmentation of Human Motion in Videos on Mobile Devices

Debski,R., Schmitt,O., Trenz,P., Reimann,M., Doellner,J., Trapp,M., Semmo,A., Pasewaldt,S.

Abstract:

This paper presents a framework and mobile video editing app for interactive artistic augmentation of human motion in videos. While creating motion effects with industry-standard software is time-intensive and requires expertise, and popular video effect apps have limited customization options, our approach enables a multitude of art-directable, highly customizable motion effects. We propose a graph-based video processing framework that uses mobile-optimized machine learning models for human segmentation and pose estimation to augment RGB video data, enabling the rendering and animation of content-adaptive graphical elements that highlight and emphasize motion.

Our modular framework architecture enables effect designers to create diverse motion effects that include body pose-based effects such as glow stick or light trail effects, silhouette-based effects such as halos and outlines, and layer-based effects that provide depth perception and enable interaction with virtual objects.

F03: Error-Robust Indoor Augmented Reality Navigation: Evaluation Criteria and a New Approach

Scheibert,O., MÃ¶ller,J., Grogorick,S., Eisemann,M.

Abstract:

Tracking errors severely impact the effectiveness of augmented reality display techniques for indoor navigation. In this work we take a look at the sources of error and accuracy of existing tracking technologies. We derive important design criteria for robust display techniques and present objective criteria. These serve evaluation of indoor navigation techniques without or in preparation of quantitative user studies. Based on these criteria we propose a new error tolerant display technique called Bending Words, where words move along the navigation path guiding the user. Bending Words outranks the other evaluated display techniques in many of the tested criteria and provides a robust, error-tolerant alternative to established augmented reality indoor navigation display techniques.

F05: Visualization of deviations between different geometries using a multi-level voxel-based representation

Dietze, A., Grimm, P., Jung, Y.

Abstract:

We present an approach for visualizing deviations between a 3d printed object and its digital twin. The corresponding 3d visualization for instance allows to highlight particularly critical sections that indicate high deviations along with corresponding annotations. Therefore, the 3d printing thus needs to be reconstructed in 3d, again. However, since the original 3d model that served as blueprint for the 3d printer typically differs topology-wise from the 3d reconstructed model, the corresponding geometries cannot simply be compared on a per-vertex basis. Thus, to be able to easily compare two topologically different geometries, we use a multi-level voxel-based representation for both data sets. Besides using different appearance properties to show deviations, a quantitative comparison of the voxel-sets based on statistical methods is added as input for the visualization. These methods are also compared to determine the best solution in terms of the shape differences and how the results differ, when comparing either voxelized volumes or hulls. The application VoxMesh integrates these concepts into an application and provides the possibility to save the results in form of voxel-sets, meshes and point clouds persistently, that can either be used by third party software or VoxMesh to efficiently reproduce and visualize the results of the shape analysis.

F11: Operational theater generation by a descriptive language

Ghiotto,M., Desbenoit,B., Raffin,R.

Abstract:

3D landscapes generation is an interdisciplinary field that requires expertise in both computer graphics and geographic informations systems (GIS). It is a complex and time-consuming process. In this paper, we present a new approach to simplify 3D environment generation process, by creating a go-between data-model containing a list of available source data and steps to use them. To feed the data-model, we introduce a formal language that describes the process"s sequence. We propose an adapted format, designed to be human-readable and machine-readable, allowing for easy creation and modification of the scenery. We demonstrate the utility of our approach by implementing a prototype system to generate 3D landscapes with a use-case fit for multipurpose simulation. Our system takes a description as input and outputs a complete 3D environment, including terrain and feature elements such as buildings created by chosen geometrical process. Experiments show that our approach reduces the time and effort required to generate a 3D environment, making it accessible to a wider range of users without extensive knowledge of GIS. In conclusion, our custom language and implementation provide a simple and effective solution to the complexity of 3D terrain generation, making it a valuable tool for users in the area.

F17: Real-Time Reflection Reduction from Glasses in Videoconferences

Tucholke,M-A., Christoph,M., Anders,L., Ochlich,R., Grogorick,S., Eisemann,M.

Abstract:

Surrounding lighting conditions cannot always be sufficiently controlled during videoconferences, yielding situations in which disturbing reflections might appear on the participants glasses. In this article, we present a retrained neural network to convincingly reduce such reflections. For real time performance we propose an asynchronous processing pipeline accompanied by a head pose-based caching strategy to reuse intermediate processing results. The implementation as virtual webcam allows the system to be used with arbitrary videoconferencing systems.

F23: The Method of Mixed States for Interactive Editing of Big Point Clouds

Benger, W. , Voicu, A., Baran, R., Gonciulea, L., Barna, C., Steinbacher, F.

Abstract:
We present a novel methodological approach for the interactive editing of big point clouds. Based on the mathematics of fiber bundles, the proposed approach to model a data structure that is efficient for visualization, modification and I/O including an unlimited multi-level set of editing states useful for expressing and maintaining multiple undo histories. Backed by HDF5 as high performance file format, this data structure naturally allows persistent storage for the history of modification actions, an unique new feature of our approach. The challenges of visually based manual editing of big point clouds are discussed and a proper rendering solution is presented. The implemented solution and its features as consequences of the underlying methodology is compared with two major mainstream applications providing point-cloud editing tools as well.

F29: First Considerations in Computing and Using Hypersurface Curvature for Energy Efficiency

Hauenstein,J., Newman,T.

Abstract:
Energy consumption for computing and using hypersurface curvature in volume dataset analysis and visualization is studied here. Base usage and usage when certain optimization steps, including compiler optimizations and variant memory layout strategies, are considered for both analysis and volume visualization tasks. Focus here is on x86, which is popular and has power measurement capabilities. The work aims to advance understanding of computing's energy footprint and to provide guidance for energy-responsible volume data analysis.

F31: Anomaly Detection with Transformer in Face Anti-spoofing

Abduh,L. , Omar,L., Ivrissimtzis,I.

Abstract:
Transformers are emerging as the new gold standard in various computer vision applications, and have already been used in face anti-spoofing demonstrating competitive performance. In this paper, we propose a network with the ViT transformer and ResNet as the backbone for anomaly detection in face anti-spoofing and compare the performance of various one-class classifiers at the end of the pipeline, such as one-class SVM, Isolation Forest, and decoders. Test results on the RA and SiW databases show the proposed approach to be competitive as an anomaly detection method for face anti-spoofing.

G02: Temporal Segmentation of Actions in Fencing Footwork Training

Malawski,F., Krupa,M.

Abstract:
Automatic analysis of actions performed in sports training can provide useful feedback for athletes. Fencing is one of the sports disciplines in which the correct technique for performing actions is very important. For any practical application, temporal segmentation of movement in continuous training is crucial. In this work, we consider detecting actions in a sequence of fencing footwork exercises. We apply pose estimation to RGB videos and then we perform per-frame motion classification, using both classical and deep learning methods. For evaluation, we provide extended manual labels for a fencing footwork dataset previously used in other works. Results indicate that the proposed methods are effective at detecting several actions. In the evaluation of our approach, we provide also a comparison with other data modalities, including depth pose estimation and inertial signals. Finally, we include an example of qualitative analysis of the performance of detected actions, to show how this approach can be used for training support.

F43: MS-PS: A Multi-Scale Network for Photometric Stereo With a New Comprehensive Training Dataset

Hardy,C., Queau,Y., Tschumperle,D.

Abstract:
The photometric stereo (PS) problem consists in reconstructing the 3D-surface of an object, thanks to a set of photographs taken under different lighting directions. In this paper, we propose a multi-scale architecture for PS which, combined with a newly designed dataset, yields state-of-the-art results. Our proposed architecture is flexible: it permits to consider a variable number of images as well as variable image size without loss of performance. In addition, we define a set of constraints to allow the generation of a relevant synthetic dataset to train convolutional neural networks for the PS problem. Our proposed dataset is much larger than pre-existing ones, and contains many objects with challenging materials having anisotropic reflectance (e.g. metals, glass). We show on publicly available benchmarks that the combination of both these contributions drastically improves the accuracy of the resulting estimated normal fields, in comparison with previous state-of-the-art methods.

F47: Fast Incremental Image Reconstruction with CNN-enhanced Poisson Interpolation

Erzar,B., Lesar,Z., Marolt,M.

Abstract:
We present a novel image reconstruction method from scattered data based on multigrid relaxation of the Poisson equation and convolutional neural networks (CNN). We first formulate the image reconstruction problem as a Poisson equation with irregular boundary conditions, then propose a fast multigrid method for solving such an equation, and finally enhance the reconstructed image with a CNN to recover the details. The method works incrementally so that additional points can be added, and the amount of points does not affect the reconstruction speed. Furthermore, the multigrid and CNN techniques ensure that the output image resolution has only minor impact on the reconstruction speed. We evaluated the method on the CompCars dataset, where it achieves up to 40% error reduction compared to a reconstruction-only approach and 9% compared to a CNN-only approach.

F67: Blocky Volume Package: a Web-friendly Volume Storage and Compression Solution

Lesar,Å½., Bohak,C., Marolt,M.

Abstract:
The Blocky Volume Package (BVP) format is a distributed, platform-independent and API-independent format for storing static and temporal volumetric data. It is designed for efficient transfer over a network by supporting sparse volumes, multiple resolutions, random access, and streaming, as well as providing a strict framework for supporting a wide palette of encoding formats. The BVP format achieves this by dividing a volume or a volume sequence into blocks that can be compressed and reused. The metadata for the blocks are stored in separate files so that a client has all the information required for loading and decoding the blocks before the actual transmission, decoding and rendering take place. This design allows for random access and parallel loading and has been specifically designed for efficient use on the web platform by adhering to the current living standards. In the paper, we compare the BVP format with some of the most often implemented volume storage formats, and show that the BVP format supports most major features of these formats while at the same time being easily implementable and extensible.

F71: Generating Realistic River Patterns with Space Colonization

Feng, H., Wuensche, B., Shaw, A.

Abstract:
River generation is an integral part of realistic terrain generation, since rivers shape terrains and changes in terrain, e.g., due to tectonic movements can change the path of rivers. Fast existing terrain generation methods often result in non-realistic river patterns, whereas physically-realistic techniques, e.g., building on erosion models, are usually slow. In this paper we investigate whether the Space Colonization Algorithm can be modified to generate realistic river patterns. We present several extensions of the Space Colonization Algorithm and show with a user study with $n=55$ participants that some variants of the algorithm are capable of generating river patterns that are indistinguishable from real river patterns. Although our technique can not generate all types of natural river patterns, our results suggest that it can prove useful for developing plausible 2D maps and potentially can form the basis for new terrain generation techniques.

F73: Using the Adaptive HistoPyramid to Enhance Performance of Surface Extraction in 3D Medical Image Visualisation

Padinjarathala,A.,Sadleir,R.

Abstract:
There are currently a range of different approaches for extracting iso-surfaces from volumetric medical image data. Of these, the HistoPyramid appears to be one of the more promising options. This is due to its use of stream compaction and expansion which facilitates extremely efficient traversal of the HistoPyramid structure. This paper introduces a novel extension to the HistoPyramid concept that entails incorporating a variable reduction between the HP layers in order to better fit volumes with arbitrary dimensions, thus saving memory and improving performance. As with the existing HistoPyramid techniques, the adaptive version lends itself to implementation on the GPU which in turn leads to further performance improvements. Ultimately, when compared against the best performing existing HistoPyramids, the adaptive approach yielded a performance improvement of up to 20\% without any impact on the accuracy of the extracted mesh.

F97: Autograsping Pose of Virtual Hand Model Using the Signed Distance Field Real-time Sampling with Fine-tuning

Puchalski M., Wozna-Szczeniak B.

Abstract:
Virtual hands have a wide range of applications, including education, medical simulation, training, animation, and gaming. In education and training, they can be used to teach complex procedures or simulate realistic scenarios. This extends to medical training and therapy to simulate real-life surgical procedures and physical rehabilitation exercises. In animation, they can be used to generate believable pre-computed or real-time hand poses and grasping animations. In games, they can be used to control virtual objects and perform actions such as shooting a gun or throwing a ball. In consumer-grade VR setups, virtual hand manipulation is usually approximated by employing controller button states, which can result in unnatural final hand positions. One solution to this problem is the use of pre-recorded hand poses or auto-grasping using physics-based collision detection. However, this approach has limitations, such as not taking into account non-convex parts of objects, and can have a significant impact on performance. In this paper, we propose a new approach that utilizes a snapshot of the Signed Distance Field (SDF) of the area below the user's hand during the grab action. By sampling this 3D matrix during the finger-bending phase, we obtain information about the distance of each finger part to the object surface. We compare our solution with those relying on physics collision detection, considering both visual results and computational impact.

G03: Accuracy of Legendre Moments for Image Representation

Bustacara-Medina,C., Ruiz-Garcia,E.

Abstract:
Existing works in orthogonal moments are mainly focuses on optimizing the classical orthogonal cartesian moments, such as Legendre moments, Gaussian-Hermite moments, Gegenbauer moments and Chebyshev moments. The current research directions generally include accurate calculation, fast calculation, robustness/invariance optimization, definition extension, and application. In this paper is presented a comparison to achieve an accuracy computation of Legendre Moments using the advances proposed by Hosny, Pawlak, Holoborodko, and a traditional technique for approximating integration computation.

G11: Optimised Light Rendering through Old Glass

Huan,Q., Rousselle,F., Renaud,C.

Abstract:

We propose a rendering method for efficiently computing the transmitted caustics produced by a glass panel with arbitrary surface deformations, characteristic of old glass used in 3D reconstructions in virtual heritage. Using Fermat's principle of least time, we generalize the concept of Next Event Estimation to allow light sampling through two displaced refractive interfaces, which amount to numerically finding all stationary points of an objective function. Our work allows for an efficient estimation of the caustic while staying inside a standard Monte Carlo pathtracing framework. Our specific geometrical context allows our solver to converge significantly faster than the more general method Specular Manifold Sampling, while scaling well with the number of panels present in the scene.

G13: Versatile Input View Selection for Efficient Immersive Video Transmission

KlÃ³ska, D., Dziembowski, A., Samelak, J.

Abstract:

In this paper we deal with the problem of the optimal selection of input views, which are transmitted within an immersive video bitstream. Due to limited bitrate and pixel rate, only a subset of input views available on the encoder side can be fully transmitted to the decoder. Remaining views are " in the simplest approach " omitted or " in the newest immersive video encoding standard (MPEG immersive video, MIV) " pruned in order to remove less important information. Selecting proper views for transmission is crucial in terms of the quality of immersive video system userâ's experience. In the paper we have analyzed which input views have to be selected for providing the best possible quality of virtual views, independently on the viewport requested by the viewer. Moreover, we have proposed an algorithm, which takes into account a non-uniform probability of userâ's viewing direction, allowing for the increase of the subjective quality of virtual navigation for omnidirectional content.

G17: Massively Parallel CPU-based Virtual View Synthesis with Atomic Z-test

Stankowski, J., Dziembowski, A.

Abstract:

In this paper we deal with the problem of real-time virtual view synthesis, which is crucial in practical immersive video systems. The majority of existing real-time view synthesizers described in literature require using dedicated hardware. In the proposed approach, the view synthesis algorithm is implemented on a CPU increasing its usability for users equipped with consumer devices such as personal computers or laptops. The novelty of the proposed algorithm is based on the atomic z-test function, which allows for parallelization of the depth reprojection step, what was not possible in previous works. The proposal was evaluated on a test set containing miscellaneous perspective and omnidirectional sequences, both in terms of quality and computational time. The results were compared to the state-of-the-art view synthesis algorithm " RVS.

G37: Evolutionary-Edge Bundling with Concatenation Process of Control Points

Saga, R., Beak, J.

Abstract:

Edge bundling is one of the information visualization techniques, which bundle the edges of a network diagram based on certain rules to increase the visibility of the network diagram and facilitate the analysis of key relationships among nodes. As one of evolutionary-based edge bundling, genetic algorithm-based edge bundling (called GABEB) is proposed which uses a genetic algorithm to optimize the placement of edges based on aesthetic criteria. However, it does not sufficiently consider the bundling between neighboring edges, and thus visual clutter issues still remain. Based on the above background, we propose an improved bundling method that considers the concatenating of control points at each edge using GABEB.

G41: Low-Rank Rational Approximation of Natural Trochoid Parameterizations

Balint, Cs., Valasek, G., Gergo, L.

Abstract:

Arc-length or natural parametrization of curves traverses the shape with unit speed, enabling uniform sampling and straightforward manipulation of functions defined on the geometry.

However, Farouki and Sakkalis proved that it is impossible to parametrize a plane or space curve as a rational polynomial of its arc-length, except for the straight line. Nonetheless, it is possible to obtain approximate natural parameterizations that are exact up to any epsilon. If the given family of curves possesses a small number of scalar degrees of freedom, this results in simple approximation formulae applicable in high-performance scenarios. To demonstrate this, we consider the problem of finding the natural parametrization of ellipses and cycloids. This requires the inversion of elliptic integrals of the second kind. To this end, we formulate a two-dimensional approximation problem based on machine-epsilon exact Chebhysev proxies for the exact solutions. We also derive approximate low-rank and low-degree rational natural parametrizations via singular value decomposition.

The resulting formulas have minimal memory and computational footprint, making them ideal for computer graphics applications.

SHORT papers

E73: Detection of Printed Circuit Board Defects with Photometric Stereo and Convolutional Neural Networks

Hable,A., Matore,M., Scherr,A., Krivec,T., Gruber,D.

Abstract:
The quality inspection of printed circuit boards (PCBs) is no longer feasible by human inspectors due to accuracy requirements and the processing volume. Automated optical inspection systems must be specifically designed to meet the various inspection requirements. A photometric stereo setup is suitable for the inspection of highly reflective gold areas on PCBs. In this process, several image acquisitions are performed under different illumination directions. This can reveal defects that are not visible under other illumination systems. In this paper, we use a segmented ring light so that image acquisition is possible under four different illumination directions. Using these images, a normal map and a mean image are calculated. The defects on the gold areas of PCBs are detectable in either the normal map, the mean image, or both images. A CNN for classification detects the defects. The input is a 6-dimensional image from normal map and mean image. An accuracy up to 95% can be achieved with the available dataset.

F19: A new deterministic gasket fractal based on ball sets

Soto-Villalobos,R., Benavides-Bravo,F.G., Hueyotl-Zahuantitla,F., Aguirre-Lopez,M.A.

Abstract:
In this paper, we propose a new gasket fractal constructed in a deterministic iterated function system (IFS) way by means of interacting ball and square sets in R^2. The gasket consists of the ball sets generated by the IFS, possessing also exact self-similarity. All this leads to a direct deduction of other properties and a clear construction methodology, including a dynamic geometry procedure with an open-source construction protocol. We also develop an extended version of the fractal in R^n. Some resulting configurations consisting of stacked 2D-fractals are plotted. We discuss about potential applications of them in some areas of science, focusing mainly on percolation models. Guidelines for future work are also provided.

F41: The Usage of the BP-Layers Stereo Matching Algorithm with the EBCA Camera Set

Kaczmarek, A.L.

Abstract:
This paper is concerned with applying a stereo matching algorithm called BP-Layers to a set of many cameras. BP Layers is designed for obtaining disparity maps from stereo cameras. The algorithm takes advantage of convolutional natural networks. This paper presents using this algorithm with a set called Equal Baseline Camera Array. This set consists of up to five cameras with one central camera and other ones aground it. Such a set has similar advantages as a stereo camera. In particular this equipment is suitable for providing 3D vision for autonomous robots operating outdoors. The research presented in this paper shows the extent to which results of using BP Layers are improving because of using the EBCA set instead of a stereo camera.

F59: Detail preserving non-rigid shape correspondences

Bindal,M., Kamat,V.

Abstract:
Understanding shapes is an organic process for us (humans) as this is fundamental to our interaction with the surrounding world. However, it is daunting for the machines. Any shape analysis task, particularly non-rigid shape correspondence is challenging due to the ever-increasing resolution of datasets available. Shape Correspondence refers to finding a mapping among various shape elements. The functional map framework deals with this problem efficiently by not processing the shapes directly but rather specifying an additional structure on each shape and then performing analysis in the spectral domain of the shapes. To determine the domain, the Laplace-Beltrami operator has been utilized generally due to its capability of capturing the global geometry of the shape. However, it tends to smoothen out high-frequency features of shape, which results in failure to capture fine details and sharp features of shape for the analysis. To capture such high-frequency sharp features of the shape, this work proposes to utilize a Hamiltonian operator with gaussian curvature as an intrinsic potential function to identify the domain. Computationally it is defined at no additional cost, keeps global structural information of the shape intact and preserves sharp details of the shape in order to compute a better point-to-point correspondence map between shapes.

F79: Using the Adaptive HistoPyramid to Enhance Performance of Surface Extraction in 3D Medical Image Visualisation

Padinjarathala,A.,Sadleir,R.

Abstract:
There are currently a range of different approaches for extracting iso-surfaces from volumetric medical image data. Of these, the HistoPyramid appears to be one of the more promising options. This is due to its use of stream compaction and expansion which facilitates extremely efficient traversal of the HistoPyramid structure. This paper introduces a novel extension to the HistoPyramid concept that entails incorporating a variable reduction between the HP layers in order to better fit volumes with arbitrary dimensions, thus saving memory and improving performance. As with the existing HistoPyramid techniques, the adaptive version lends itself to implementation on the GPU which in turn leads to further performance improvements. Ultimately, when compared against the best performing existing HistoPyramids, the adaptive approach yielded a performance improvement of up to 20 percent without any impact on the accuracy of the extracted mesh.

F89: Semi-Supervised Learning Approach for Fine Grained Human Hand Action Recognition in Industrial Assembly

Sturm, F., Sathiyababu, R., Hergenroether, E., Siegel, M.

Abstract:
Until now, it has been impossible to imagine industrial manual assembly without humans due to their flexibility and adaptability. But the assembly process does not always benefit from human intervention. The error-proneness of the assembler due to disturbance, distraction or inattention requires intelligent support of the employee and is ideally suited for deep learning approaches due to the permanently occurring and repetitive data patterns. However, there is the problem that the labels of the data are not always sufficiently available. In this work, a spatio-temporal transformer model approach is used to address the circumstances of few labels in an industrial setting. A pseudo-labeling method from the field of semi-supervised transfer learning is applied for model training, and the entire architecture is adapted to the fine-grained recognition of human hand actions in assembly. This implementation significantly improves the generalization of the model during the training process over different variations of strong and weak classes from the ground truth, proving that it is possible to work with deep learning technologies in an industrial setting even with few labels. In addition to the main goal of improving the generalization capabilities of the model by using less data during training and exploring different variations of appropriate ground truth and new classes, the spatial recognition capabilities of the model are improved by adding convolution to the spatial embedding layer, which increases the test accuracy by over 5% compared to a similar predecessor model.

G05: Position Based Rigid Body Simulation: A comparison of physics simulators for games

Seabra,M., Fernandes,F., Lopes,D., Pereira,J.

Abstract:
Interactive real-time rigid body simulation is a crucial tool in any modern game engine or 3D authoring tool. The quest for fast, robust and accurate simulations is ever evolving. PBRBD (Position Based Rigid Body Dynamics), a recent expansion of PBD (Position Based Dynamics), is a novel approach for this issue.
This work aims at providing a comprehensible state-of-the art comparison between Position Based methods and other real-time simulation methods used for rigid body dynamics using a custom 3D physics engine for benchmarking and comparing PBRBD (Position Based Rigid Body Dynamics), and some variants, with state-of-the-art simulators commonly used in the gaming industry, PhysX and Havok. Showing that PBRBD can produce simulations that are accurate and stable, excelling at maintaining consistent energy levels, and allowing for a variety of constraints, but is limited in its handling of stable stacks of rigid bodies due to the propagation of rotational error.

G29: On Importance of Scene Structure for Hardware-Accelerated Ray Tracing

Kacerik,M., Bittner,J.

Abstract:
Ray tracing is typically accelerated by organizing the scene geometry into an acceleration data structure. Hardware-accelerated ray tracing, available through modern graphics APIs, exposes an interface to the acceleration structure builder that builds it given the input scene geometry. This builder is, however, opaque, with limited control over the internal building algorithm. Additional control is available through the layout of AS builder input data, a geometry of the scene structured in a user-defined way. In this work, we evaluate the impact of a different scene structuring on the performance of the ray-triangle intersections in the context of hardware-accelerated ray tracing. We discuss the possible causes of significantly different outcomes for the same scene and outline a solution in the form of automatic input structure optimization.

POSTERS

D37: On Unguided Automatic Colorization of Monochrome Images

Sluzek, A.

Abstract:
Image colorization is an ill-posed problem with an infinite number of RGB solutions for a single grayscale picture. Thus, human assistance either direct (e.g. scribbling) or indirect (e.g. by providing analogous images, or large training sets of sufficiently representative images) is considered indispensable to perform colorization in a visually plausible manner. Therefore, the objective of this paper looks slightly seditious (if not impossible) as we strive to perform image colorization without any additional metadata or human hints, using a grayscale image as the only source of data. We just assume an rgb-to-gray model converting color images (possibly not even physically existing) into the grayscale ones, and use a few straightforward heuristics. The results are surprisingly repeatable (in spite of strong presence of probabilistic elements in the proposed approach) and visually acceptable, which suggests that colorized insights into 'gray worlds" (even if only for aesthetic reasons) are feasible. Thus, the approach is intended mainly for domains where only monochrome visual representations physically exist (and convincingly rich color visualizations should be synthesized). The paper contains description of the method, experimental results and selected supplementary topics (e.g. a quest for relevant quality metrics).

E03: Justice expectations related to the use of CNNs to identify CSAM. Technological interview.

Oronowicz-Jaskowiak, W., Wasilewski, P., Kowaluk, M.

Abstract:
A technological interview was conducted with representatives of the judiciary in order to determine their expectations and beliefs related to the technological solution (involving detection of child sexual abuse materials using CNNs), being developed. The obtained results lead to the following conclusions: 1. Representatives of the judiciary recognize the advantages of the technological solution being created in the form of accelerating the work of experts and minimizing the risk of mistakes. 2. Representatives of the judiciary see the limitations of the technological solution being created in the form of the inability to replace court experts and emphasize that it also depends on the stage of the case. 3.The selection of pornographic materials from a specific set for later verification by a forensic expert is of the greatest importance.

E11: The use of artificial intelligence for automatic waste segregation in the garbage recycling process

Bobulski,J., Kubanek,M.

Abstract:
The problem of recycling secondary raw materials remains unresolved, despite many years of work on this issue. Among the many obstacles that arise is also the difficulty of sorting individual waste fractions. To facilitate this task and help solve this problem, modern computer vision and artificial intelligence techniques can be used. In our work, we propose constructing an intelligent garbage bin containing a camera and a microcomputer along with software that uses these techniques to sort waste. The role of the software is to recognize the type of waste and assign it to one of five main categories: paper, plastic, metal, glass and cardboard. The proposed method uses image recognition techniques with a convolutional neural network. The results confirm that using artificial intelligence methods significantly helps in sorting waste. The proposed solution can be used in homes and public places such as parks, cinemas or playgrounds.

F02: Detection of Dangerous Situations Near Pedestrian Crossings using In-Car Camera

Kubanek, M., Karbowiak, L., Bobulski, J.

Abstract:
The paper presents a method for detecting dangerous situations near pedestrian crossings using an in-car camera system. The approach utilizes deep learning-based object detection to identify pedestrians and vehicles, analyzing their behavior to identify potential hazards. The system incorporates vehicle sensor data for enhanced accuracy. Evaluation results show high accuracy in detecting dangerous situations. The proposed system can potentially enhance pedestrian and driver safety in urban transportation.

F13: Autonomous Photogrammetry Workflow for Real Time Rendering using Free Software

Pardo,R., Remolar,I.

Abstract:
Real-time applications require low-poly 3D objects to ensure the interactivity and a high frame rate. Free 3D models available on Internet usually are composed by a vast amount of polygons and this represents a problem for interactive applications. Photogrammetry can be a solution to create 3D objects, but their complexity and adaptation to the real-time applications is also a problem. Some applications have appeared that decimate the geometry, but they usually require be managed by experimented 3D artists or do not give the better results without paying a fee. In order to produce appropriate low-poly models, this work presents a work-flow that allows inexperienced designers create realistic low poly models using free software. Using the proposed solution, indie developers will be able to create their own realistic 3D assets with cheap resources and only using available free software for all platforms. This solution will make it possible to save time and costs.

F83: Compressed Deep Learning Model for Detecting COVID-19 Disease: A Genetic Algorithm based approach

Reddy,S., Reddy,G., Gupta,S., Hazra.D., Gupta,U.

Abstract:
Data processing and analysis are becoming more difficult because of the development of Internet of Things (IoT) devices and cameras, which has resulted in increased data generation. Deep learning, a branch of artificial intelligence, has become a critical tool for processing these gadgetsâ' massive volumes of data. As a result, developments like intelligent houses, driverless vehicles, and surveillance systems have been created. However, the complexity of deep learning models and the data required for training can lead to high computational costs and large model sizes. This has led to the need for model compression techniques to reduce the size of deep learning models without sacrificing performance. However, substantial computing costs and big model sizes might occur from the complexity of deep learning models and the volume of data needed for training. Using compression methods, these models can be made smaller without compromising accuracy. In conclusion, the growth of IoT devices and cameras has increased the demand for deep learning methods of data analysis. Deep learning models can be more effective using model reduction, making them better suited for Internet of Things (IoT) devices with constrained storage capacity and processing.

G59: Polychromatism of all light waves: new approach to the analysis of the physical and perceptive color aspects

Niewiadomska-Kaplar,J.

Abstract:
Research on light vision mechanisms in biosystems and on the mechanisms of formation of deficits in color discrimination[1] reveals that not only white light is polychromatic but all light waves are. The spectrum of white light is composed of aggregations of only 4 monochromatic waves: magenta UV 384 nm, cyan 432 nm, yellow 576 nm and magenta IR 768 nm, grouped in 5 bi-chromatic waves: cinnabar red (magenta IR + yellow), green (yellow + cyan), indigo (cyan + magenta UV) and also two semi-bright bi-chromatic waves - porphyry IR (semi-infrared wave composed of the magenta IR 768 nm wave and the colorless infrared wave 864 nm) and porphyry UV (semi-ultraviolet wave composed of the magenta UV 384 nm wave and the colorless ultraviolet wave 288 nm). The light waves thus composed create the light sensations due to the mechanism of additive synthesis.
The method allows a new approach to interpret the composition of the bright waves, the phenomenon of decomposition of colours and additive synthesis that constitutes the principle of colour production in computers. The new elaborate models of colour physics also constitute the basis by interpretation of the mechanisms of vision of colours.