

Guest Editorial: Machine Vision Applications

Yasuyo Kita¹ · Hiroshi Ishikawa² · Takeshi Masuda¹

Published online: 13 February 2017

© The Author(s) 2017. This article is published with open access at Springerlink.com

Machine vision, system-oriented and application-oriented subarea of computer vision, has drastically progressed over the last few decades along with the progress of computer vision theories and been playing the key roles in our daily life. In addition, the computational environment, such as massive progress of computational power and the accumulation of common big data, as well as the ubiquity of 2D and 3D cameras, enables machine vision systems to deal with a wide range of real world problems. Along this line of progress, we have decided to edit this special issue, machine vision applications, by collecting ten representative papers in the field in order to show the current and future directions in the field to the IJCV audience.

Parts/system inspection has a long history in the field and has largely progressed by boosting adaptability to complex situation. “Domain Adaptation for Automatic OLED Panel Defect Detection Using Adaptive Support Vector Data Description” (doi:[10.1007/s11263-016-0953-y](https://doi.org/10.1007/s11263-016-0953-y)) by Sindagi et al. addresses the real world problem that is a degradation of the classifier performance due to changes in inspection circumstances such as lighting configurations. The authors extend the Support Vector Data Description (SVDD) for adaptively learn an incremental classifier based on a source

classifier. They also propose a new feature descriptor using modified Local Binary Pattern (LBP) and local inlier-outlier ratios for detection of OLED panel defects. Experimental results have shown its superiority, especially in micro defects detection.

We sleep about one-third of our life, and sleep ergonomics is one of important factors for maintaining our quality of life. “Automatic Sleep System Recommendation by Multi-modal RGB-Depth-Pressure Anthropometric Analysis” (doi:[10.1007/s11263-016-0919-0](https://doi.org/10.1007/s11263-016-0919-0)) by Jordi et al. proposes a system for generating a bed mattress prescription by sensing the sleeping human body with a RGB-D sensor (Kinect) and a 2-D pressure sensor. From the measured RGB-D information, the user’s body shape parameters (weight, height, BMI, morphotype category) are estimated, and by taking the pressure information and clinical knowledge into account, the most suitable sleep system (mattress-topper-pillow) is recommended by the system. Experiments support usability of the proposed method in real stores.

Hyperspectral imaging, due to its rich information, is a popular imaging technique in many machine vision application areas including agriculture and biomedicine. The difficulty in handling hyperspectral images, however, exist in high dimensionality. “Adaptive Spatial-Spectral Dictionary Learning for Hyperspectral Image Restoration” (doi:[10.1007/s11263-016-0921-6](https://doi.org/10.1007/s11263-016-0921-6)) by Fu e al. proposes a novel hyperspectral image (HSI) restoration method that effectively utilizes underlying characteristics of HSIs. The method adaptively learns spatial-spectral dictionary through considering “high correlation across spectra” and “non-local self-similarity over space” in the degraded HIS. Then, an HSI restoration model is designed based on the local and non-local sparsity of the HSI under the learned spatial-spectral dictionary. Experimental results show the effectiveness of the proposed method for denoising and superresolution.

✉ Yasuyo Kita
y.kita@aist.go.jp
Hiroshi Ishikawa
hfs@waseda.jp
Takeshi Masuda
t.masuda@aist.go.jp

¹ National Institute of Advanced Industrial Science and Technology, AIST Tsukuba Central 1, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8560, Japan

² Department of Computer Science and Engineering, Waseda University, Okubo 3-4-1, Shinjuku, Tokyo 169-8555, Japan

Machine vision systems for automatic driving has been actively studying. In particular, the introduction of CNNs to this field is one of the key issues to realize automatic driving due to its robustness by modeling high dimensionality. “A Practical and Highly Optimized Convolutional Neural Network for Classifying Traffic Signs in Real-Time” (doi:[10.1007/s11263-016-0955-9](https://doi.org/10.1007/s11263-016-0955-9)) by Aghdam et al. describes an extensive study on how to design CNNs for traffic sign classification. The authors present an accurate, efficient, and compact practical pipeline for recognizing traffic signs in real-time, using CNNs and recent advances in deep learning. They also propose new methods for creating an ensemble of networks to analyze their stability, including traditional classification and end-to-end learning methods for solving this problem.

Recognition of human activity in daily life is one of the most popular areas in the field. “Robust Statistical Frontalization of Human and Animal Faces” (doi:[10.1007/s11263-016-0920-7](https://doi.org/10.1007/s11263-016-0920-7)) by Sagonas et al. proposes a method of normalization of face images with variations of pose, illumination and occlusion, which is efficient for face recognition and fine-grained categorization. The key idea is the frontal face image has the minimum nuclear norm, and the problem is formulated by the low-rank minimization problem with help of frontal face subspace. Experimental results show the robustness of the proposed method that can be applicable to distorted face images, face sketches, and animal face images.

“Growing Regression Tree Forests by Classification for Continuous Object Pose Estimation” (doi:[10.1007/s11263-016-0942-1](https://doi.org/10.1007/s11263-016-0942-1)) by Hara et al. presents an adaptive node splitting method for training regression forest which is referred to as the K-clusters Regression Forest (KRF). The number of clusters is determined adaptively based on the sample distribution. The proposed algorithm is applied to estimation of head pose, car direction, and pedestrian orientation, and the superior performance was obtained compared to existing regression methods in the experiments.

“Multi-Camera Multi-Target Tracking with Space-Time-View Hyper-graph” (doi:[10.1007/s11263-016-0943-0](https://doi.org/10.1007/s11263-016-0943-0)) by Wen et al. proposes a method for tracking multiple targets from surveillance videos captured from different viewpoints. The problem is solved by searching the optimal subgraph in the space-time-view hyper-graph that encodes higher-order constraints on 3D geometry, appearance, motion continuity, and trajectory smoothness within different views. Experimental results show that the proposed method performs favorably against the state-of-the-art methods with efficient computational performance.

Human activities are composed of several actions. In “Complex Activity Recognition via Attribute Dynamics” (doi:[10.1007/s11263-016-0918-1](https://doi.org/10.1007/s11263-016-0918-1)) by Li et al., the seg-

mented actions are represented by the bags of attribute sequence, and temporal attribute dynamics is learned to describe the human activities using the binary dynamic system (BDS) representation. The attribute dynamics are encoded by manifold embedding in the bag-of-words for attribute dynamics (BoWAD) and the vectors of locally aggregated descriptors (VLAD). The proposed methods are experimented on several real human activity videos with the state-of-the-art level performance.

Precise motion analysis of human and other non-rigid objects also becomes close to real applications. “Combining Local-Physical and Global-Statistical Models for Sequential Deformable Shape from Motion” (doi:[10.1007/s11263-016-0972-8](https://doi.org/10.1007/s11263-016-0972-8)) by Agudo et al. proposes a practical non-rigid structure from motion algorithm that runs sequentially. The authors introduce local motion constraints based on Newton’s second law and combine with global statistical constraints that are progressively learned in a bundle adjustment framework to reconstruct camera motion and object non-rigid motion in 3D. Their system shows similar performance to computationally intensive batch approaches in a wide range of applications.

“3D Human Pose Tracking Priors using Geodesic Mixture Models” (doi:[10.1007/s11263-016-0941-2](https://doi.org/10.1007/s11263-016-0941-2)) by Simo-Serra et al. proposes a method of approximating the probability density function (PDF) of a potentially large dataset that lies on a known Riemannian manifold. The proposed method uses a finite mixture model on the manifold, and is computationally efficient even with large datasets. This model is applied to 3D human pose tracking whose motion model is represented by a set of angles of a tree of connected joints. Experiments show that the proposed method greatly outperforms the standard Gaussian diffusion prior. The method is tested with various artificial manifolds, and due to its generality, it can be employed for many applications.

In summary, machine vision technology is drastically improving its flexibility to work in unconstrained natural scenes with the rapid advancements of computer vision and closely related fields. This issue covers from the traditional inspection techniques for industrial applications through active autonomous driving and human activity recognition for our daily life. We hope this issue to provide the readers new and wider perspectives on the research activities in machine vision application area.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.