STAG: Smart Tools & Apps for Graphics (2015), pp. 1–2
Silvia Biasotti and Marco Tarini (Editors)

Multimedia Technology and Telecommunications Lab

Giancarlo Calvagno, Pietro Zanuttigh, Simone Milani, Fabio Dominio, Giampaolo Pagnutti, Giulio Marin, Ludovico Minto

University of Padova

Abstract

The Multimedia Technology and Telecommunications Laboratory of the Department of Information Engineering of the University of Padova has a long time research record in 3D data acquisition, processing and transmission. Current research areas include 3D data acquisition with a particular focus on consumer depth cameras, hand gesture recognition from color and depth data, compression and transmission of image, video and 3D data, multimedia forensics, super resolution of depth and color data, scene segmentation and content-based retrieval of colored 3D objects.

This document presents the staff of the Laboratory and then briefly introduces the main research areas.

1. Staff

Currently there are 3 permanent staff members, 3 PhD students and one student with a research grant.

1.1. Giancarlo Calvagno (Associate Professor)

Giancarlo Calvagno was born in 1962. He received the engineering degree in 1986 and the PhD degree in 1990, both from the University of Padova. From 1988 to 1990 he was at the University of Illinois at Urbana-Champaign as a Visiting Scholar. From 1990 to 1998 he was with the Department of Information Engineering of the University of Padova as an assistant professor. Since 1998 he has been an Associate Profesor in the same Department. He has been involved in several national and European research projects on image and video coding and processing.

1.2. Pietro Zanuttigh (Assistant professor)

Pietro Zanuttigh was born in 1978. He graduated in Computer Engineering at the University of Padova in 2003 and got the PhD from same university in 2007. In 2007 he got the assistant professor position at the University of Padova. He has also been a visiting researcher at the University of New South Wales (Sydney, Australia) in 2006 and in 2010. Now he works in the Multimedia Technology and Telecommunications group and his research activity is focused on 3D data processing. His research interest includes the transmission and remote visualization of 3D scenes, the acquisition of 3D data with a particular focus on sensor fusion, the compression of depth maps and multi-view videos and gesture recognition.

1.3. Simone Milani (Assistant professor)

Simone Milani received his Laurea Degree from the University of Padova, Italy in 2002. In 2004 he was enrolled in the Ph.D. course in Electronics and Telecommunication Engineering at the Department of Information Engineering (DEI) of the University of Padova. His research was focused on video source coding and video transmission over lossy networks. From 2011 to 2014 he has been working at the Politecnico di Milano, Italy within the EU project REWIND. Now he is a researcher at the University of Padova. His main research topics are digital signal processing, source coding, joint source-channel coding, robust video transmission over lossy packet networks, distributed source coding, 3D video coding and processing, multimedia forensics.

1.4. Other members

- Giampaolo Pagnutti (PhD Student)
- Ludovico Minto (PhD Student)
- Giulio Marin (PhD Student)
- Fabio Dominio (Research Grant)

2. Research Areas

2.1. ToF (Time-of-Flight) and stereo data fusion

This research work considers heterogeneous acquisition systems made of standard cameras together with depth (Time-Of-Flight) cameras. We derived a probabilistic fusion algorithm that allow us to obtain high quality depth information from the data of both the ToF camera and the stereo-pair. The fusion algorithm is derived in a probabilistic setup that allows the decoupling of the information from the stereo pair and the information from the ToF camera. Accurate models for the measurement errors of the stereo and ToF systems are derived and then used into the probabilistic fusion framework. The more advanced version of the approach [DZC15] exploits a measurement error model accounting for the mixed pixels effect together with a global MAP-MRF optimization scheme using an extended version of Loopy Belief Propagation with site-dependent labels.

2.2. Compression of image video and 3D data

Research activities involving data compression concern the compression of depth data and 3D video representations. In the recent work in collaboration with the University of New South Wales (AU) the use of a scalable breakpoint field together with a breakpoint adaptive Wavelet decomposition is proposed for the coding of the depth field [MTZ13]. We also exploited the information about the objects in the scene to perform a cognitive 3D compression of both color and depth streams. Depth maps contain key information on the scene structure that can be effectively exploited to improve the performance of multi-view coding schemes. We introduced a novel coding architecture that replaces the inter-view motion prediction operation with a 3D warping approach based on depth information to improve the coding performances. The pixels of the different warped views are packed into a stack of aligned views which can be efficiently coded by transform coding techniques.

2.3. Hand Gesture Recognition with color and depth data

Depth data acquired by current low-cost real-time depth cameras provide a more informative description of the hand pose that can be exploited for gesture recognition purposes. Following this rationale, we proposed a novel hand gesture recognition scheme based on depth information. The basic framework [DDZ14] is the following: color and depth data are firstly used together to extract the hand and divide it into palm and finger regions. Then different sets of feature descriptors are extracted accounting for different clues like the distances of the fingertips from the hand center, the curvature of the hand contour or the geometry of the palm region. Finally a multi-class SVM classifier is employed to recognize the performed gestures.

2.4. Handheld scanning with depth cameras

The exploitation of consumer depth cameras for 3D scanning purposes is a very challenging task due to the limited accuracy and reliability of their data. We proposed a 3D reconstruction pipeline explicitly targeted to the Kinect. The proposed scheme aims at obtaining a reliable reconstruction that is not affected by the limiting issues of these cameras and is at the same time simple and fast in order to allow using the Kinect sensor as an hand-held scanner. A novel algorithm for the extraction of salient points exploiting both depth and color data is firstly proposed. Then the extracted points are used within a modified version of the ICP algorithm that exploits both geometry and color distances.

2.5. Scene segmentation from color and depth data

We introduced a novel segmentation scheme where multi-dimensional vectors are used to jointly represent color and depth data, and various clustering techniques are applied to them in order to segment the scene [DZC12]. Furthermore we addressed the critical issue of how to balance the two sources of information by using an automatic procedure based on an unsupervised metric for the segmentation quality. Different acquisition setups, like Time-of-Flight cameras, the Microsoft Kinect device and stereo vision systems have been used inside this framework. We also proposed an improved segmentation scheme that exploits a 3D surface estimation scheme.

2.6. Other research areas

- Multimedia forensics
- Super resolution of depth and color data
- Transmission and remote visualization of 3D scenes
- Muon Tomography
- Markerless motion capture
- Content-based retrieval of colored 3D objects

References

[DDZ14] DOMINIO F., DONADEO M., ZANUTTIGH P.: Combining multiple depth-based descriptors for hand gesture recognition. *Pattern Recognition Letters* (2014), 101–111. 2

[DZC12] DAL MUTTO C., ZANUTTIGH P., CORTELAZZO G. M.: Fusion of geometry and color information for scene segmentation. *IEEE Journal of Selected Topics in Signal Processing* (September 2012). 2

[DZC15] DAL MUTTO C., ZANUTTIGH P., CORTELAZZO G. M.: Probabilistic tof and stereo data fusion based on mixed pixels measurement models. *IEEE Transaction on Pattern Analysis and Machine Intelligence* (2015). 2

[MTZ13] MATHEW R., TAUBMAN D., ZANUTTIGH P.: Scalable coding of depth maps with rd optimized embedding. *IEEE Transactions on Image Processing* 22, 5 (2013), 1982–1995. 2