INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 04 MPEG VIDEO CODING

ISO/IEC JTC 1/SC 29/WG 04 m66990

April 2024, Rennes, France

Title	**[MIV] A new natural content proposal: Choreo**
Source	**PUT, ETRI**
Authors	**Dominika Klóska, Dawid Mieloch, Adrian Dziembowski, Błażej Szydełko, Jakub Stankowski, Gwangsoon Lee**

Abstract

This document presents a proposal of new natural content that was captured by a system of 20 cameras arranged in an arch. The proposed content has two versions that introduce different levels of challenges for processing. For the purpose of this document, the authors focused on the version that introduces a lesser amount of challenges for the encoding and decoding process. Future work will be focused on obtaining depth maps of better quality using separate textures for still objects and moving objects.

1 Proposal

The camera arrangement for this sequence consisted of 20 cameras arranged on an arc. Cameras were placed on 5 camera rigs, and each rig contained 4 equally-distributed cameras, where the distance between neighboring devices was set to 25 centimeters.

Two versions of the sequence were captured – one with the curtains closed and one with curtains opened in order to introduce different challenges for processing. The comparison is presented below:

View 0

View 9

View 18

Intrinsic paramerers of each camera were estimated, views captured by cameras were undistorted:

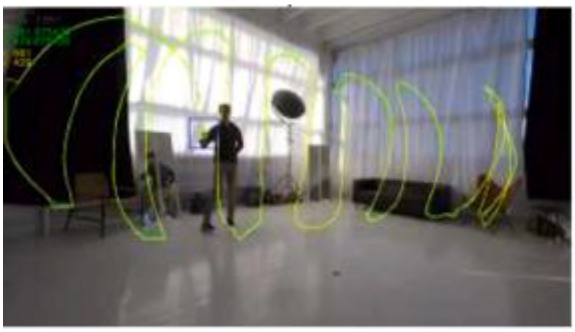
View 9 – raw:

View 9 – undistorted:

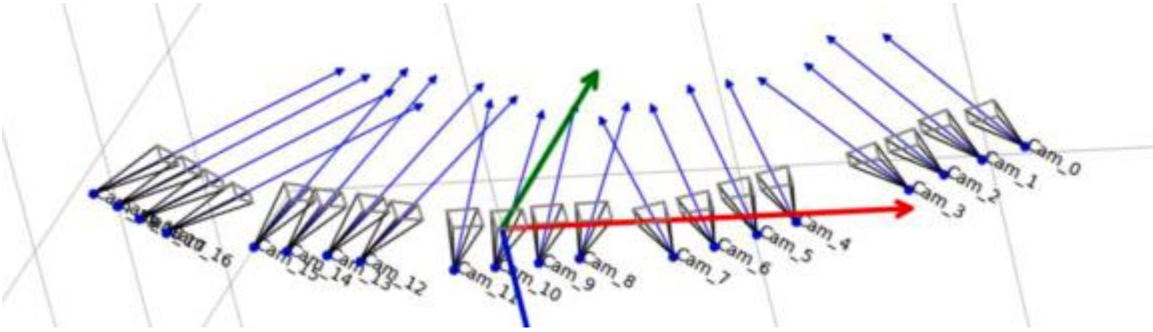
In order to estimate extrinsic parameters of the multicamera system, the calibration sequence containing spherical orange marker was captured. The sequence was processed within the ECPC framework:

A single frame of calibration sequence

Tracked marker

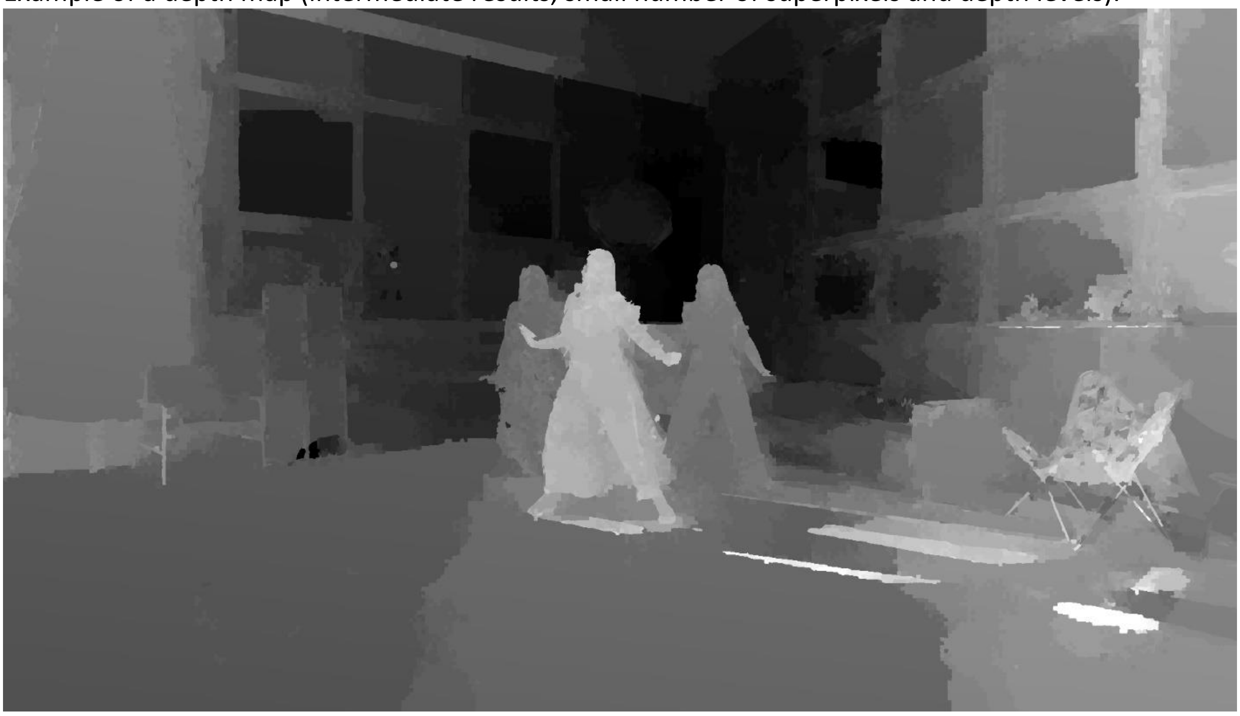


Visualization of multicamera system based on estimated extrinsic parameters



Using input views, camera parameters, and IVDE software, the depth maps can be calculated:

Example of a depth map (intermediate results, small number of superpixels and depth levels):



The posetrace created from these textures and depth maps is presented below:

Frame 0

Frame 31

Frame 198

In order to obtain depth maps of better quality authors decided to divide the textures into two categories: textures containing only still objects and textures containing only moving objects. In order to extract moving objects from the sequence the detectron software was used. Because of challenges introduced by this content the software produced output that had to be corrected in order to guarantee better quality of depth maps. The software seemed to label further objects as moving when they where dynamically occluded by the front, actually moving objects. Apart from this scenario the software also labeled still objects as moving in random single frames throughout the sequence. These problems were addressed by introducing some slight modifications in the code as well as modifying the software parameters separately for each view.

Detectron software results BEFORE modifications

Detectron software results AFTER modifications

This concludes the process of extracting moving objects into separate textures.

As for still objects textures they were obtained by calculating median over time of pixels in the calibration sequence. The result can be seen below:

View 0

View 9

View 18

Future work will be focused on estimating depth maps of abovementioned textures as well as creating posetraces with the use of new depth maps.

2. Acknowledgement

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2018-0-00207, Immersive Media Research Laboratory).