1. Public Executive Summary

The task 3D Image Capture combines tri-focal stereovision, spatio-temporal structured light and a cinematographic camera with Z-channel into a unified video-and-depth recording device. The device will consist of a central cinematographic camera, two or four small satellite cameras mounted to the left and to the right hand side of the central camera and a beamer for projecting structured light, mounted on top or below the central camera.

All devices will be mounted on a rig. Due to vibrations, the relative positions of two cameras to each other may deviate with six degrees of freedom; three translational and three rotational. Assuming that only HD (1920x1080) cameras with a normal lens (focal length = 2500 pixel) are used and defining that deviations due to vibrations should not exceed the size of 1 pixel the maximal translational deviation is 0.8mm between focal points along the baseline as well as perpendicular to the baseline and 2.1mm between focal points along the optical axis. The maximum rotational deviation is 0.02° for pan and tilt around the baseline and 0.06° for roll around the optical axis.

Meeting these requirements a rig was designed and produced. The result allows to place a satellite camera on the rig in less than 5 minutes and to change the position of a satellite camera in less than 1 minute. The rig weights less than 24kg.

In the first project phase the tri-focal and the structured light approach will be optimized independently. The tri-focal system uses Hitachi HV-F31CL cameras with XGA resolution and Camera Link interface as satellites. The signals are grabbed using standard PCs and Matrox Helios XCL boards. The black level of the CVBS output of the central camera is converted to TTL and used as trigger input for the grabber boards which then distribute the trigger to the satellite cameras. The sequences from all cameras have to be pair wise rectified and saved in single image sequences with bmp or png format. Results are stored in a proprietary format providing the depth per pixel in mm along with 2 consistency measures.

The structured light system uses Point Grey grasshopper firewire-B cameras at full HD resolution, but at a reduced frame rate, which is sufficient for research purposes. In cases a higher sensitivity is needed, prosilica 1.4 megapixel cameras are used, with GigE vision interface. For projecting patterns, either a modified data beamer is used, or a custom designed LED-based multi-slide projector. The projector and cameras are synchronised with the VIPER by means of a custom designed hardware box developed jointly by GVN and UoH in the context of the RACINE-IP project. The sequences from the machine vision cameras are stored in a proprietary format developed at UoH, aimed at gaining maximum performance on off-the-shelf PC and laptop platforms. Camera-projector and camera-camera point correspondence maps are output, again in a custom developed format. These data is further processed into depth maps with confidence.

In the second project phase the approaches will be merged towards a unified system. The partners will monitor the market for industry cameras to find small HD cameras with good image quality and GigE or HD-SDI interface. HD-SDI would have to be converted to GigE. The 10GigE signal from the central camera and the 1GigE signals from the satellite cameras are sent to a switch which is connected to a Field Recorder. Input and output sequences will use DPX format. The format for the calibration information will be developed in collaboration with WP3 and documented in D3.1.

For more information about this documents, please contact: info@20203dmedia.eu