CMOS Vision Chip for Gaussian Filtering

M. Suárez[1], V.M. Brea[1], J. Fernández-Berni[2,3], R. Carmona-Galán[2], D. Cabello[1], A. Rodríguez-Vázquez[2,3]

[1]Centro de Investigación en Tecnoloxías da Información (CITIUS), University of Santiago de Compostela, Santiago de Compostela, Spain

[2]CSIC, Instituto de Microelectrónica de Sevilla (IMSE-CNM), Seville, Spain

[3]University of Seville, Instituto de Microelectrónica de Sevilla (IMSE-CNM), Seville, Spain

Email: victor.brea@usc.es

Feature detectors are a category of algorithms in computer vision intended for applications such as visual tracking, object recognition or segmentation [1]. Invariance against illumination levels, partial occlusions or geometrical transformations as rotations or scaling are key issues in the design of a high quality feature detector. The invariance against scaling in feature detectors like Scale Invariant Feature Transform (SIFT) is achieved with the so-called Gaussian pyramid [2]. The Gaussian pyramid comprises a set of octaves. Each octave is an image of half resolution of the former octave. In turn, every octave contains a set of scales. The scales are built up by filtering with Gaussian filters of increasing width or sigma levels. The construction of the Gaussian pyramid demands high computational time. A dedicated CMOS chip which incorporates image acquisition is a good approach to reach fast and power-efficient image analysis. This demo will include real-time tests of a proof-of-concept CMOS vision chip on standard 180 nm technology that performs the Gaussian pyramid. The chip captures an image of 176 x 120 px. It also performs the Gaussian pyramid with a switched-capacitor network in a double-Euler configuration. The area of the chip amounts to 5 x 5 mm2. The photosensors are nwell photodiodes in a 3-Transistor Active Pixel Configuration (3T-APS) occupying 7.4 x 6.7 μm2. Groups of 4 photodiodes are arranged in Processing Elements (PEs). The PEs contain local circuitry for analog to digital conversion and correlated double sampling. The chip provides the Gaussian pyramid of 3 octaves and 6 scales each with an energy cost of 26.5 nJ at 2.64 Mpx/s with an RMSE when compared to a pure software solution running on a PC below 1.2% of full- scale value.

[1] Tinne Tuytelaars and Krystian Mikolajczyk, “Local Invariant Feature Detectors: A Survey”, Foundations and Trends in Computer Graphics and Vision, vol. 3, no. 3 (2007) 177-280.

[2] D.G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”. Int. J. of Comp. Vision 60(2), pp. 91-110, 2004.