Moving object tracking on panoramic images

Péter Földesy[*+], István Szatmári[*+], Ákos Zarándy[*+]

[*]AnaLogic Computers Ltd. 13-17 Kende Street, Budapest, h-1111 Hungary
[+]Computer and Automation Research Institute of the Hungarian Academy of Sciences (MTA-SZTAKI), 13-17 Kende Street, Budapest, h-1111 Hungary
email: fpeter@analogic-computers.com

Abstract. Analogic CNN algorithms are combined with the special properties of a panoramic imaging device called PAL optics. The resulting system is capable to identify and track moving objects in 360 degrees field of view real-time. Image processing engine was the ALADDIN system with a 64x64 array size CNNUM chip. As the image sensor, we tested the optical input of the CNNUM chip in one experiment, and in other cases, we used a regular camera.

1 Introduction

Security and surveillance applications often need the permanent monitoring of 360 degrees view area. Traditionally this problem is solved by either with a rotated camera or by a number of cameras with slightly overlapping view areas. Both solutions have a number of disadvantages. Just to mention the technical problems, the rotated camera solution raises problems including the non-static images and the non-permanent field coverage. The multiple camera solution leads to boundary problems when a tracked moving object is moving from one camera field to another. On the other hand, both solution is expensive, because either the rotating platform or the multiple cameras and optics, and frame grabbers are needed. In this paper, we introduce an elegant, cost effective way of solving this problem.

We combined two promising technologies. The first one is the Panoramic Annular Lens (PAL) system [1], which projects the 360 degrees field of view into a single disc-shaped image (**Fig. 1.**). The second technology is the Aladdin Visual Computer [2,3], which is powered by the ACE4K 64x64 array size CNNUM chip [4]. The combination of the panoramic imaging device and the large computational power of the CNN technology [5,6] leads to a real-time, 360 degrees view angle, single camera monitoring system. The system and its algorithms are described in this paper. Moreover, we present here two different experimental results. In the first case, we used the on-chip optical sensor of the ACE4K chip, while in the second case, we used a regular camera, which provided much higher resolution.

2 The Aladdin Visual Computer

The Aladdin Visual Computer is the high performance, standalone computational environment of the ACE4K chip. The system is implemented on PC-104plus form factor boards. In this way, standard industrial quality PC-104plus PC hosts a specially designed DSP module, which drives the platform board carrying the ACE4K chip (**Fig. 1.**). The other role of the high performance DSP is to compute those operations which cannot be executed on the ACE4K efficiently. To reach high-speed communication, PCI bus is used as a communication channel between the host and the DSP module. The DSP module, the platform, and the hosting industrial PC constitute a rugged, high performance unit what we call Aladdin visual computer. The Aladdin Visual Computer can be plugged into dsktop PC also. More information about the system can be found in [2,3].

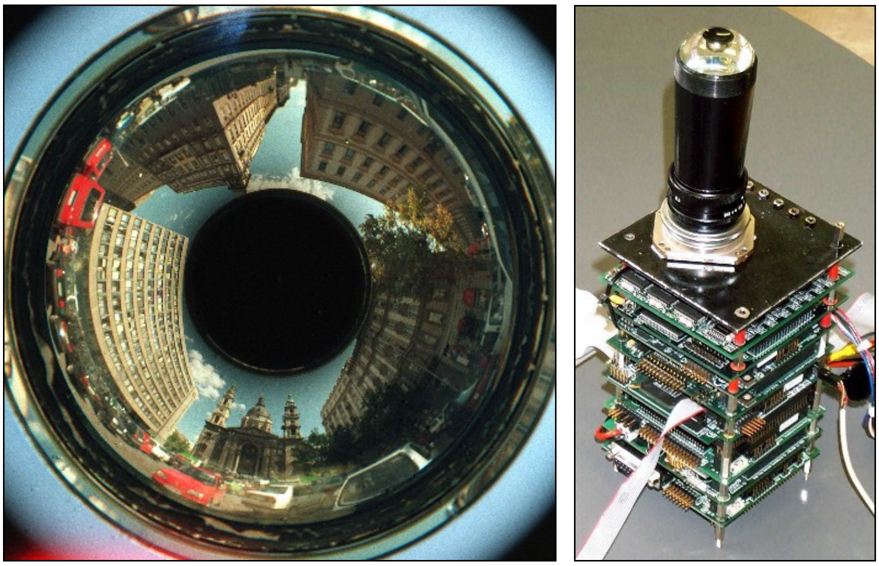


Fig. 1. A typical image captured by a camera with PAL optics (left). It is a 360 degrees view of a parking lot surrounded with houses. (Cars and houses behind the camera are up side down and projected to the top part of the image.) The naked Aladdin Visual Computer with the PAL optics on it (right). The PAL optics is projecting images directly to the surface of the CNNUM.

3 Algorithms and experimental results

Two groups of tests were done. The first one used the direct optical input of the ACE4K chip through the PAL optics. The second test series was made on image sequences prerecorded by a camera using PAL optics. These two groups were separated because the image quality of a panoramic image on a low resolution chip is quite poor, but still some features can be extracted, while it the image quality is much better if a regular camera is used. The targeted problem was to detect movements in the 360-degree field of view, moreover some basic classification tasks were concerned also.

3.1 Experiments using the on-chip optical input of the ACE4K with the PAL optics

The ACE4K chip has no dedicated direct optical input. Nevertheless, the CMOS surface of the analog image memories (LAMs) gives the possibility of image acquisition. The snapshots in **Fig. 2**. show images captured by the direct optical input of the ACE4K chip through the PAL optics. Due to the low light level of the PAL optics the images are quite dark in spite of long integration time.

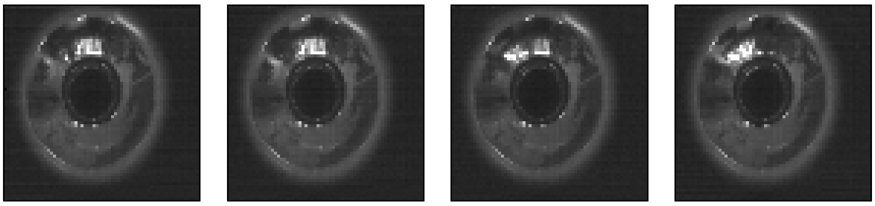


Fig. 2. Consecutive snapshots of direct optical input of the ACE4K chip using PAL optics.

The low resolution makes almost impossible to recognize objects by humans either. Since this chip was not designed to have direct optical input the light sensitivity of the chip is quite low. It means that without a strong light source this optical input cannot be used. Additional noise effects have strong influence on the images. We put efforts to filter out regular and stochastic errors. Nevertheless, some basic movement detection was made on this sequence. This is shown in **Fig. 3.**

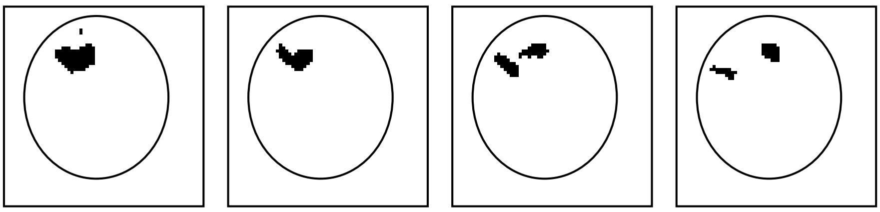


Fig. 3. Movement detection on the ACE4K chip using direct optical input with PAL optics

The movement detection algorithm is based on mainly temporal homotopic filtering, gray-scale range selection, and mathematical morphology operations. These operations were executed mainly on the ACE4K chip. The measured executing time of this algorithm proves that even fast objects can be tracked with this system.

3.2 Experiments using a camera with PAL optics

In these experiments, the optical input of the chip was not used. The PAL optics was mounted on a standard video camera and the captured images were transferred to the ACE4K chip electronically. This makes possible to test processing capabilities of the ACE4K chip without using its image capturing capabilities.

Two consecutive snapshots in **Fig. 4** show images captured by the camera through the PAL optics, meanwhile **Fig. 5** shows a transformed image for demonstration purpose. (Note, that this kind of time consuming image transformation is not a part of the algorithm.) The processed image size is 320x300. The reason why the image size is not N×N is because we used a non-square pixel camera. However, this did not cause any problem later on.

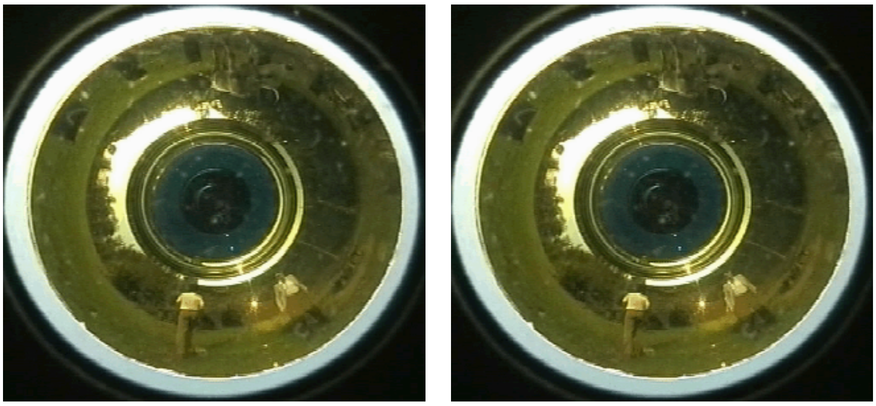


Fig. 4. Two consecutive snapshots of a video flow using PAL optics optical input.

Fig. 5. A snapshot of a video flow using PAL optics optical input converted to a stripe image for demonstration purpose.

The images were separated to 64x64 pixel sized blocks and these blocks were processed one after the other on the ACE4K chip. The cutting and merging ability is incorporated in the system in an automatic way so image sequences composed of arbitrary sized images can be processed with this system. The speed of the ACE4K chip makes possible to work on large images real-time.

The special projection of the PAL optics makes this problem far from trivial. Depending on the distance from the camera, relative sizes of objects change dramatically from small patches to almost half size of image area. Therefore, sophisticated methods are necessary to distinguish different objects and extract important features of them. The flowchart of the movement detection and classification algorithm is shown in **Fig. 6**. The motion detection and object classification algorithm has five major parts:

1. *Optical acquisition and background estimation*
2. *Motion map generation with morphological filtering*

3. *Pre-filtering characteristic feature extraction*
4. *Feature based classification and position determination*
5. *Object tracking and path tracing*

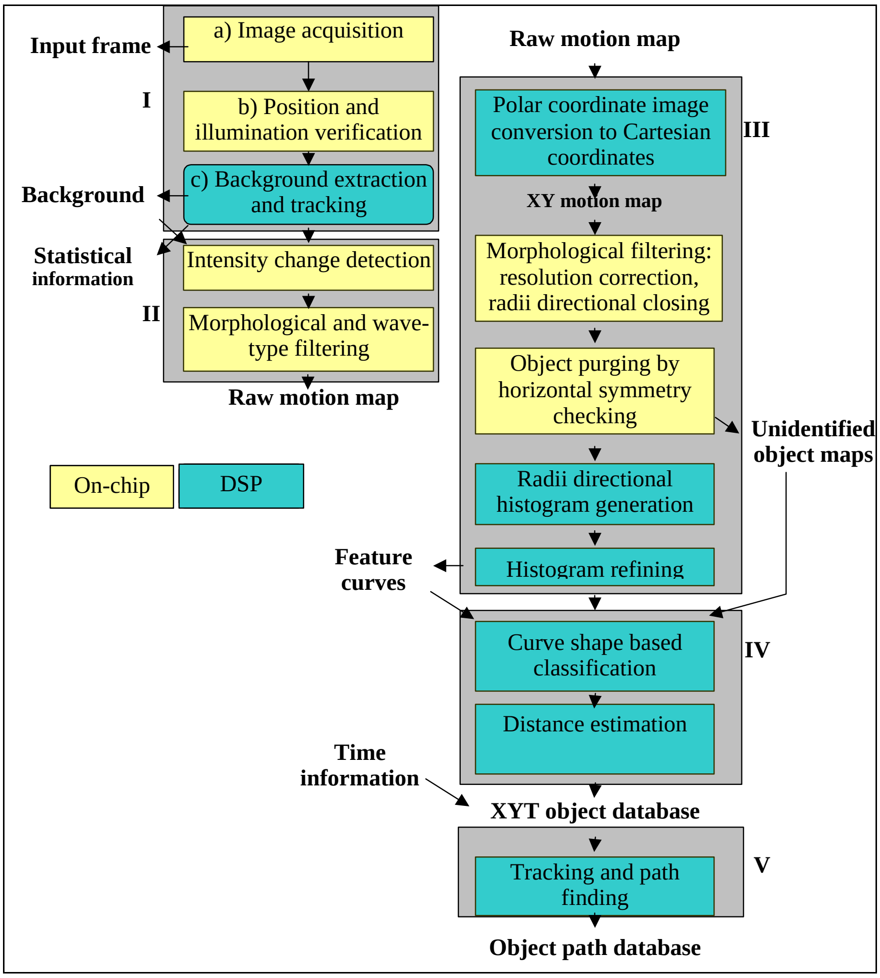


Fig. 6. Flowchart of the motion detection and object classification algorithm.

The first part of the algorithm performs background estimation continuously from the captured image sequence. This also includes an extensive pre-filtering, and temporal filtering. The second part of the algorithm generates a raw motion map. This part includes spatial morphological and wave type filtering running on the ACE4K chip. These types of operations would require too long execution time on a DSP therefore the ACE4K chip is an ideal tool to accomplish these tasks. The third part of the algorithm makes a coordinate transformation from the polar representation in order to extract characteristic features of objects. The chip also plays important role here. Statistical features are extracted on the DSP. The fourth part is the classification and position determination. It generates several true and false object candidates. The fifth part filters and revises the classification result based on dynamical class models and constructs tracking routes.

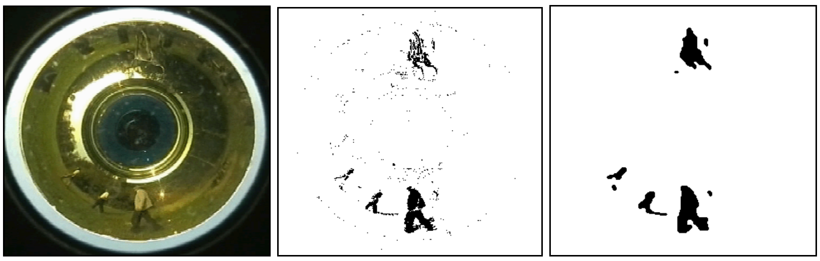
Time consumption of the movement and classification algorithm			
	Function	**Time**	**Remark**
I.	Image acquisition	50 μs	on chip image sensing
	Position and illumination verification	50μs	$3*T_{template}+5*T_{logic}$
	Background extraction and tracking (with image transfer)	410μs	with DSP
II.	Intensity change detection	20μs	
	Morphological and wave-type filtering	60μs	
III.	Polar-Cartesian transformation	80μs	
	Resolution correction and radii directional closing	50μs	with image transfer
	Object purging by horizontal symmetry checking	50μs	with image transfer
	Radii directional histogram generation	5μs	
	Histogram refining	5μs	
IV.	Curve shape based classification	100μs	
V.	Tracking and path finding	100μs	
Sub total 1 (ACE4K chip):		280μs	
Sub total 2 (DSP):		700μs	Texas, C6202 @ 250MHz
Total time 1:		980μs	chip size image
Total time 2:		40ms	large image handling (320x300)

Table 1. Measured execution time of the individual steps of the algorithm.

As it can be seen from **Table 1**, the CNNUM chip with a proper direct optical input could solve this task with roughly 1000 frame/sec while using camera with large images video rate (25Hz) can be achieved. The chip execution time is valid for not only the ACE4K chip but also larger array size chip because template and logic operations require almost the same execution time. The execution time of the DSP related operations depend strongly on the chosen DSP architecture, clock frequency, and the image size.

3.3 Experimental results

In this section, we introduce our results with the PAL camera images. We show snapshots from this sequence, and spatial-temporal 3D graphics, which includes all the moving objects tracked in time. **Fig. 7a** shows a snapshot of the original image sequence. This is followed by the moving part detection and the noise filtering (**Fig. 7b,c**). From this, the typical shapes of the humans are picked up, and the resulting positions are placed to a 3D coordinate system, where the x and y coordinates are corresponding to the location of the humans, while the z axis is representing the time. **Fig. 8** shows the raw and the filtered tracking data. **Fig. 9** shows a snapshot of the final result image sequence, on what the objects are located and marked, and their distance is shown.



a) b) c)

Fig. 7. Moving object extraction. a) captured image; b) extracted moving parts; c) morphologically filtered moving regions.

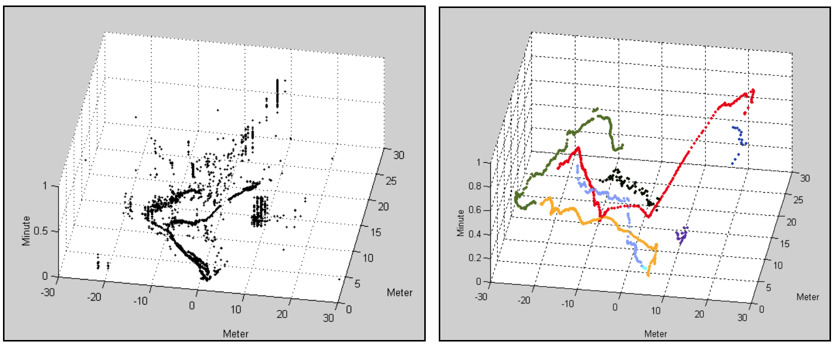


Fig. 8. Raw (left) and noise filtered (right) tracking data.

4 Conclusions

We showed that the ACE4K chip with a companion DSP is capable to solve object classification and tracking problem real-time. Due to the lack of a good direct optical input of the ACE4K chip, we had to use an external camera with PAL optics as the input source. Time speed of the proposed algorithm makes possible to track even very fast objects in the scene, although further optimization and careful system design is necessary. We believe that the ACE16K chip with array size 128x128 and with direct optical input will overcome this problem.

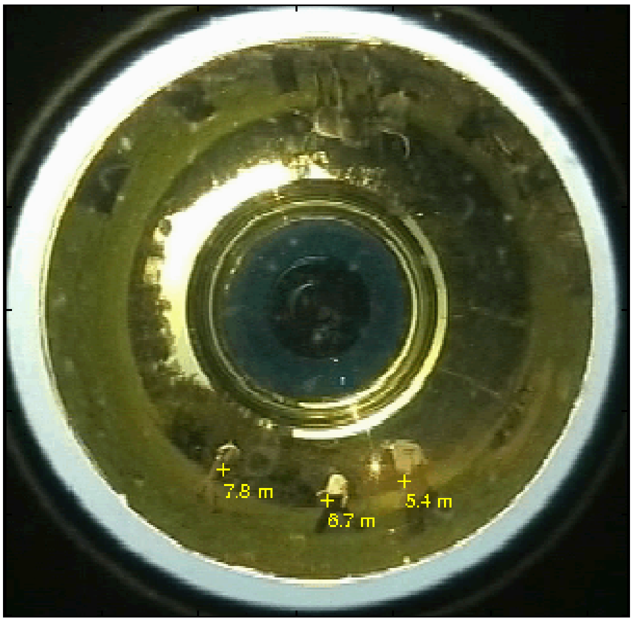


Fig. 9. The tracking information – both distance and position – is superimposed into a snapshot of the input sequence.

5 Acknowledgement

This work was partially supported by the EOARD grant No. F61775-01-WE010.

References

1. Hungarian Patent 192125, USA Patent 4566763, Japan Patent 1962784, German Patent 3402847.
2. A. Zarándy, "ACE Box: High-performance Visual Computer based on the ACE4K Analogic Array Processor Chip", European Conference on Circuit Theory and Design (ECCTD' 01), Helsinki, 2001
3. Web page of AnaLogic Computers Ltd. www.analogic-computers.com
4. S. Espejo, R. Domínguez-Castro, G. Liñán, and Á. RodriguezüVázquez, "A 64x64 CNN universal chip with analog and digital I/O" Proc. 5th Int. Conf. on Electroonics, Circuits and Systems (ICECS-98), Lisbon, Portugal, pp. 203-206 1998.
5. L.O. Chua and T. Roska, "The CNN Paradigm", *IEEE Transactions on Circuits and Systems - I*, vol. 40, no. 3, March 1993, pp. 147-156
6. T.Roska and L.O.Chua, "The CNN Universal Machine: an analogic array computer", IEEE Transactions on Circuits and Systems-II Vol.40, pp. 163-173, March,1993.