Implementation of Motion Detectors: A Case Study

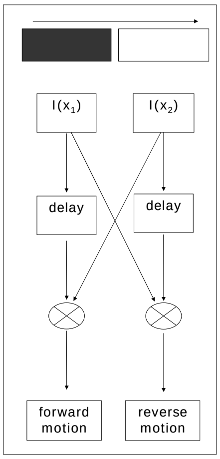
Theresa Cooke, Christian Wallraven

Introduction

- Two types of motion detectors were implemented in order to evaluate their applicability in a real-time computer vision scenario. In particular, we are interested in implementing the algorithms on an FPGA, which allows for flexible hardware implementation of algorithms at the price of introducing constraints on real-time capability.
- The Reichardt correlation dectector requires the least amount of computational effort, due to its simple architecture. However, its output is noisy and brittle. Two variants were thus implemented to improve performance, one based on averaging inputs and another based on cross-correlation. Pooling results over space and time was also implemented.
- The second type of detector implemented is a basic version of the oriented spatio-temporal energy filtering model (Adelson and Bergen).
- Detectors were evaluated both on controlled artificial stimuli and - more importantly - on real-world video sequences of moving faces.

Motion Detector Models

Reichardt Correlation Detectors



This luminance-based detector has been widely used in models of insect vision and human perception (Reichardt). The intensity values of two pixels separated in space and time are multiplied in order to generate a directionally-oriented response. The output from one mirror-symmetric unit is subtracted from the output of another so that motion can be detected in both forward (right/down) and reverse (left/up) directions. One set of detectors is implemented for horizontal motion detection and another is implemented for vertical motion detection.

Two strategies aimed to increase the stability of the Reichardt detector were implemented:

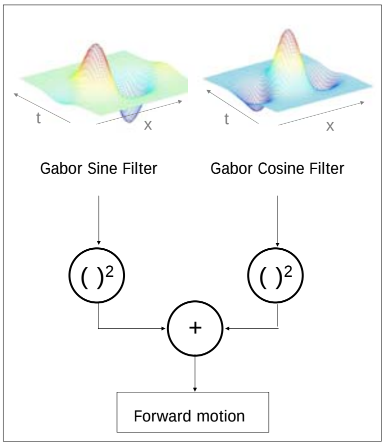
- **"Averaged Reichardt Detector"**: The single pixel input was replaced with an average of local intensity values.
- **"Normalized Cross-Correlation (NCC) Reichardt Detector"**: The multiplication is replaced by a normalized cross-correlation between two neighbourhoods of intensity values. This modification removes the Reichardt detector's sensitivity to average intensity (DC gain).

Spatio-Temporal Energy Detector

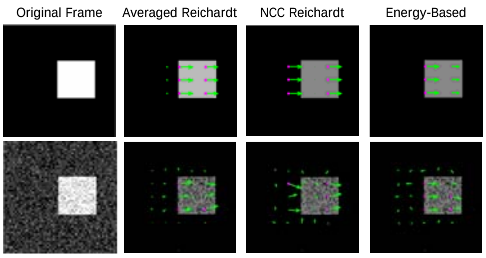
Here, motion is seen as energy with a specific orientation in space-time (Adelson). This energy can be extracted using a filter with a specific orientation in space-time. Commonly used filters are the Gabor sine and cosine filters, which extract odd and even energy, respectively. For motion detection, the image sequence is cut along the x-t or y-t plane and filtered with the Gabor pair.

The filters' output is summed and squared to provide an estimate of motion in one direction. The same output from a mirror-symmetric set of filters is subtracted in order to account for motion in the reverse direction.

One set of detectors is implemented for horizontal motion detection and another set is implemented for vertical motion detection.

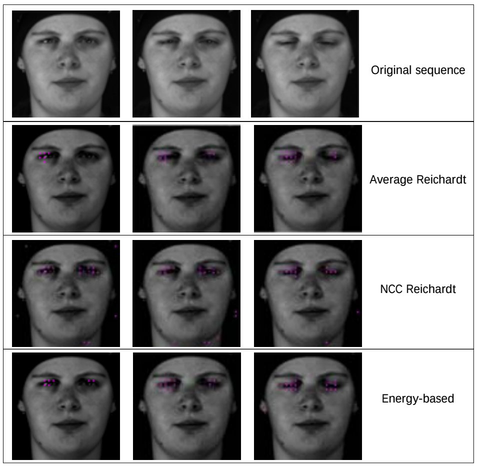


Experiment 1: Artificial Stimuli



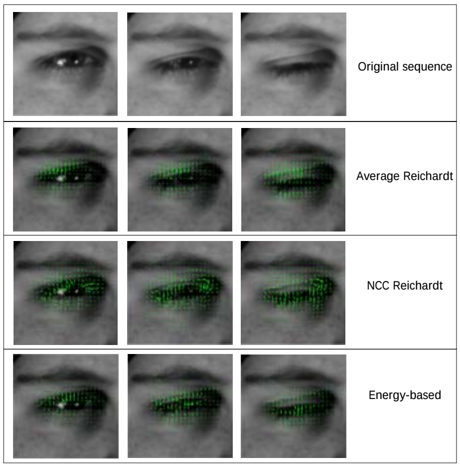
- The detectors were first tested on artificial stimuli consisting of a box moving over a uniform background at fixed velocity. Two cases, shown left, were tested: **with** and **without Gaussian noise**.
  [The direction of the arrows is determined by the amount of motion detected in horizontal and vertical directions. Dots are placed wherever the sum of squared horizontal and vertical motion exceeds 70% of the maximum value found per frame.]
- All three detectors are **sensitive to noise** introduced in the second case. The NCC Reichardt detector is particularly affected because it is insensitive to the contrast difference between the box and the background. Both the averaged Reichardt detector and the energy-based detector are contrast sensitive and are thus less susceptible to noise when there is a large contrast difference between the object and the background.
- To improve robustness in subsequent experiments, the output of a single detector was replaced by the **average** output of four directly adjacent detectors. In addition, the two Reichardt variants were **pooled over time** (the energy-based detector receives input over time due to the size of its filters.)

Experiment 2: Face Stimuli



- Here, detectors are tested on a video of a moving face.
- Dots mark locations where a strong motion signal was found.
- Movement in the left eye is favoured by both the two contrast-sensitive detectors (averaged Reichardt and energy-based). The contrast-insensitive NCC Reichardt picks up movement in the background, even though the pixel intensity values are very low.
- **The three detectors perform about equally well**, suggesting that the most computationally efficient one may be used to obtain coarse information about where motion occurs.

Experiment 3: Close-Up Stimuli



- How well can the three detectors estimate the direction of motion?
- Here, a small region of the full-resolution video around the left eye is fed into the detector bank.
- All three detectors are reasonably accurate, but the NCC Reichardt provides more information about the direction of motion. For instance, it is able to capture motion in parts of the eye cast in shadow.
- The other two detectors are **distracted** by movement of high intensity points such as light reflections.

Conclusions

- The detectors' performance is affected by contrast and noise. The averaged Reichardt and the energy-based detectors prefer high contrast areas, which is an advantage when the high-contrast area is of interest. The NCC Reichardt is sensitive to noise regardless of the average intensity. This makes it more sensitive to motion in shadowed areas, but also more sensitive to background noise. Thus a contrast-sensitive method may be preferable for coarse motion detection, while a contrast-invariant method could be preferable for fine analysis of motion direction.
- A clear advantage of the Reichardt variants is the simplicity of their base architecture, which can be fine-tuned in order to achieve the desired trade-off between accuracy and computational complexity. For instance, more stable results can be achieved with pooling over space and time. The two variants tested here provided performance comparable to that of the energy-based detector, which is more costly computationally.
- Given these results, a possible **architecture for motion analysis** is coarse motion detection using the averaged Reichardt detector on the FPGA, where computational complexity must be minimized. Regions of interest could be signaled upstream/offline, where a finer analysis of motion direction could be performed using the NCC Reichardt or the energy-based detectors.

References

Adelson, E. H. & Bergen, J.R. (1985). Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A. 2, 284-299.

Mallot, H.A. (2000) Computational Vision: Information Processing in Perception and Visual Behaviour. MIT Press Cambridge 178-199.

Reichardt W, (1959) Autocorrelation and the central nervous system. In W. A. Rosenblith (Ed.) Sensory Communication MIT Press Cambridge 303-318.

Zanker JM, Hofmann MI, Zeil J (1997) A two-dimensional motion detector model (2DMD) responding to artificial and natural image sequences. Invest Ophthalmol Vis Sci 38: S 936.