Adaptive Motion Pooling and Diffusion for Optical Flow

N V Kartheek Medathati*, Manuela Chessa,
Guillaume S. Masson, Pierre Kornprobst, Fabio Solari

Summary

We study the impact of local context of an image (contrast and 2D structure) on spatial motion integration by MT neurons. To do so, we revisited the seminal work by Heeger and Simoncelli (HS) [4] using spatio-temporal filters to estimate optical flow from V1-MT feedforward interactions. However, the HS model has difficulties to deal with several problems encountered in real scenes (e.g., blank wall problem and motion discontinuities). Here, we propose to extend the HS model with adaptive processing by focussing on the role of local context indicative of the local velocity estimates reliability. We set a network structure representative of V1, V2 and MT areas of the motion stream. We incorporate three functional principles observed in primate visual system: contrast adaptation [3], adaptive afferent pooling [2] and MT diffusion that are adaptive dependent upon the 2D image structure (Adaptive Motion Pooling and Diffusion, AMPD). We evaluated both HS and AMPD models performance on Middlebury optical flow estimation dataset [1]. Our results show that the AMPD model performs better than the HS model and its overall performance is comparable with many modern computer vision methods. The AMPD model could be further improved by integrating feedback to better recover true velocities around motion boundaries. We propose that such adaptive model can serve as a ground for future research in biologically-inspired computer vision.

Model At a given time, V1 activity is modelled by classical motion energy $E^{V1}(p,\theta)$ where p is spatial position and θ is the preferred direction of the contrast sensitivity in the spatial domain. V2 activity is represented by $R(p)$ representing local image structure through oriented Gabor functions. MT activity is obtained by an adaptive pooling defined by:

$$E^{MT}(p,d) = F\left(\sum_{i=1}^{N} cos(\theta_i - d)\frac{1}{N(p,\theta_i)}\sum_{p'} W_R(p,p')E^{V1}(p',\theta_i)\right),$$

where F is a non-linear function, N is a normalization and spatial pooling is regulated in terms of spatial extension and anisotropy through $W_R(p,p')$: Pooling becomes more local and anisotropic close to image discontinuities (see examples in Fig. 1). Adaptive diffusion is defined by an iterative procedure starting with regions of high confidence defined by a cornerness measure based on texture energies.

Results A sample result on a sequence from Middlebury dataset is shown in Fig. 1. Results show an improvement in the estimation near boundaries.

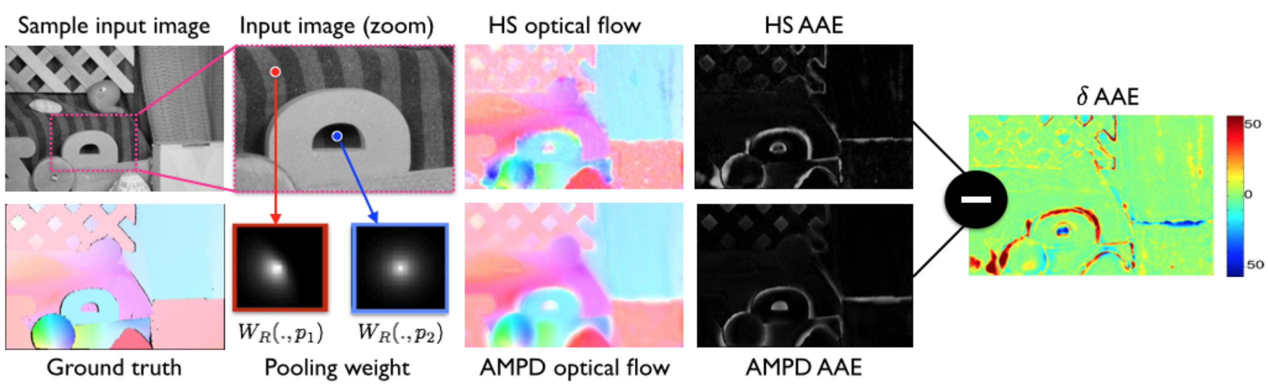

Figure 1: Sample results using sequence "RubberWhale" from Middlebury training set. Optical flow are represented using the color from [1]. We show how pooling weight are adaptive depending on position and comparisons of optical flow between HS and APMD with Average Angular Error (AAE).

References

[1] S Baker, et. al., A database and evaluation methodology for optical flow. *IJCV*, 2011.

[2] X. Huang, et. al., Adaptive surround modulation in cortical area MT. *Neuron*, 2007.

[3] M. P. Sceniak, et. al., Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neuroscience*, 1999.

[4] E. P. Simoncelli et. al., A model of neuronal responses in visual area MT. *Vision Research*, 1998.

*KM and PK acknowledge funding from the EC IP project FP7-ICT-2011-8 no. 318723 (MatheMACS)