Evaluating Motion Estimation Models from Behavioural and Psychophysical Data

Émilien Tlapale[1], Pierre Kornprobst[1], Jan D. Bouecke[2], Heiko Neumann[2], and Guillaume S. Masson[3]

[1] Neuromathcomp Lab, INRIA Sophia Antipolis – Méditerranée
http://www-sop.inria.fr/neuromathcomp
[2] Vision and Perception Science Lab, University of Ulm
http://www.informatik.uni-ulm.de/ni/mitarbeiter/HNeumann/
[3] DyVA Lab, INCM, UMR 6193 CNRS – Université de la Méditerranée
http://www.incm.cnrs-mrs.fr/equipedyva.php

Abstract. Offering proper evaluation methodology is essential to continue progress in modelling the neural mechanisms involved in vision information processing. Currently the evaluation of biologically inspired motion estimation models lacks a proper methodology for comparing their performance against behavioural and psychophysical data. Here we set the basis for such a new benchmark methodology based on human visual performance and designed a database of image sequences taken from neuroscience and psychophysics literature. In this article we focused on two fundamental aspects of motion estimation, which are the respective influence between 1D versus 2D cues and the dynamics of motion integration. Since motion models deal with many kinds of motion representations and scales, we defined two general readouts based on a global motion estimation. Such readouts, namely eye movements and perceived motion, will serve as a reference to compare simulated and experimental data. Baseline results are provided for biologically inspired artificial vision models but also for computer vision models. As a whole we provide here the basis for a valuable evaluation methodology to unravel the fundamental mechanisms of motion perception in the visual cortex. Our database is freely available on the web together with scoring instructions and results at: http://www-sop.inria.fr/neuromathcomp/psymotionbench

Keywords: evaluation methodology, biologically inspired artificial vision models, motion estimation, optical flow, motion perception, eye movements.

1 Introduction

Offering proper evaluation methodology is essential to continue progress in modelling the neural mechanisms involved in vision information processing. This general idea has been very well understood and applied in computer vision where

J. Suzuki and T. Nakano (Eds.): BIONETICS 2010, LNICST 87, pp. 483–496, 2012.
© Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2012

challenging benchmarks are now available for several key problems allowing models to be compared and further improved. For example, motion estimation performance in computer vision increased significantly thanks to several classical benchmarks, which pointed out strength and weaknesses of state of the art approaches [3, 2].

The benchmark for optical flow introduced by Baker and co workers [2] defines a set of challenging image sequences with associated ground truth. The choice of sequences was guided by the needs to evaluate models performance on key difficulties encountered by modellers (motion at objects boundaries, occlusions, non-rigid motions, large displacements). The proposed evaluation methodology consisted of several quantified criteria based on local comparisons between ground truth and output from computer vision models.

Herein we define a motion evaluation methodology where the visual system performance acts as ground truth. Since models aim at elucidating both the computational principles and the computing architectures involved in motion processing, comparing their outputs to biological responses is therefore a strong requirement. Such an evaluation methodology is very different from classical computer vision benchmarks where flow fields are compared together. In our context, the notion of local motion does not make a lot of sense when considering the visual system performance since the purpose of the visual system is not to estimate a dense flow field. Thus defining global readouts is necessary in order to compare output from models with observable quantities measured in neuroscience experiments. Moreover behavioural and perceptual experiments provide numerous types of data such as perceived motion direction and speed or smooth pursuit eye movements.

In this paper we set the basis for a new benchmark methodology which is based on human visual performance. Section 2 describes the main difficulties to design such an evaluation methodology. Section 3 proposes several stimuli and associated readouts selected for motion estimation evaluation. Section 4 presents baseline results. The scoring procedure is illustrated on an example and we show our baseline results which includes both biologically inspired artificial vision models and computer vision models. Section 5 concludes and mentions possible extensions of this work.

Stimuli, scoring procedure and baseline results are available online at:
http://www-sop.inria.fr/neuromathcomp/psymotionbench

2 Comparison Difficulties

2.1 Stimulus Parametrisation

Comparing models performance to biological data requires the definition of an homogeneous stimuli set. For example the stimuli are characterised by their physical size, the distance to the observer, and their visual field size. In a benchmark stimuli set we need to ensure a constant mapping between the physical and the

numerical dimensions, otherwise incoherences will occur in the simulated results. Other physical quantities such as duration and luminance also require precise mappings.

2.2 Discretisation

Stimuli discretisation implies the necessity to define properly a scale factor for converting real-world values characteristics into computer parameters. The main problem with discretisation procedures is the *aliasing problem*. One has to make sure that frequency of the input does not cross the Nyquist frequency.

The spatio-temporal discretisation maps the time and the visual field as a succession of discrete images uniformly sample Such a representation is geometrically different from the log-polar retinotopic disposition of the visual cortex. Moreover the precision of the input has to be sufficient to avoid aliasing problems with the stimuli sizes used by the experimentalists. The finer the precision the larger the data and a good compromise between those two quantities has to be chosen.

Luminance is usually encoded by an eight bits values at each pixel. Such a coarse quantisation is a severe restriction since contrast has a profound impact of the temporal dynamics of most visual percepts and is responsible for many dynamical non-linearities.

2.3 Inhomogeneity of the Motion Representations

All motion models do not have the same motion representation. Their output can be described by global velocity likelihoods [25], velocity distributions at every position [15, 5], filter responses [1], time-correlated spike trains [13], or 2D flow fields [3]. A typical biological model of motion integration might include V1 layers with filter-like responses, MT layers corresponding to local pattern translational motions and MST layers giving indications of global rotation or expansion motions. Because of the variety of motion representations it is necessary to define common observable quantities which are comparable to experimental measurements. These common observable quantities are called *readouts* in both simulations and experiments presented herein.

2.4 Lack of Ground Truth

In computer vision the ground truth is the true velocity field, which is easily defined for synthetic videos, and which can also be estimated for real scene videos. For example in [2], the authors proposed videos of real scenes with the true velocity field. Algorithms can be evaluated based on local comparisons of the 2D flow fields against the estimated flow resulting from different algorithms.

In psychophysical studies the notion of ground truth is less obvious and it is impossible to define it in a strict sense. For example, one has to handle the great variability between subjects or between trials for a single subject. The concision

of data reported in the literature, often a mean and a standard deviation, does not allow the extraction of the statistical laws underlying the data. Moreover many stimuli are bi-stable or multi-stable, and an additional difficulty in defining a ground truth.

Among the set of experimental stimuli studied in neuroscience some provided results at different levels. For instance, the coherence level necessary for perceiving global motion in a random dot patterns has been measured in human subjects but also in single neurons in areas MT, MST, and LIP. In the 2D motion integration example being considered here, a consistent set of global direction estimates have been collected at these different levels as well as for human perception, and monkey and human smooth pursuit [22]. When available these datasets collected for different responses with a single set of motion stimuli should be used to benchmark models.

Given the diversity of the neuroscience experiments, capturing the main properties and results of motion estimation appears to be a complex task. For this reason we restrict our study to a set of fundamental questions described in the following section.

3 Database Design

In this paper we focus on two fundamental aspects of motion integration. Namely, we want to evaluate models performance with stimuli showing the respective influence between 1D versus 2D cues, and the dynamics of motion integration. We chose four stimuli fitting into two classes: line-drawings objects and gratings (see Figure 1). For the purpose of our evaluation we selected stimuli for which smooth pursuit eye movements and motion perception data were available. This section presents the stimuli in more details.

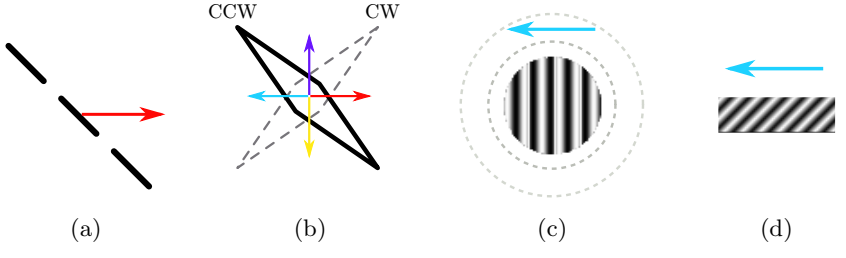


Fig. 1. Database design. The proposed stimuli fit into two classes: line drawings and gratings. (a) Translating bar. (b) Translating diamond. (c) Grating size. (d) Barber pole.

3.1 Line-Drawing Objects

Translating Bars. In [20, 8] the authors consider tilted translating bars. Pursuing a translating bar which true motion direction is not normal to its orientation leads to initial deviation in the *smooth pursuit* eye movement direction. For example, the peak directional error for a tilted bar of 20° length, 0.5° width and 92 cd/m^2 luminance, moving with 16°/s is about 30° [8].

To obtain a model evaluation procedure, the slope of the directional errors could be analysed with respect to bar length, number of bar tiles. Indeed, as the bar length is increased it becomes more complicated to recover its true direction. Likewise it easier to pursue one long bar, if it is tiled into several sub bars [20, 8].

Translating Diamonds. In [24] the authors consider diamond stimuli translating either vertically or horizontally. Due to the local orientations of the diamonds edges with respect to the translating direction, these stimuli mimic type II plaids. Indeed the vector average of the edge motions is biased 44° away from the object's direction. The stimuli thus provide an interesting example to study the influence of 1D and 2D cues on motion integration.

Changing the configuration of the stimulus, by using clockwise (CW) or counterclockwise (CCW) stimuli, or by varying the direction of the translation, does not influence the ability to pursuit the translating diamonds. In all the cases, the initial pursuit direction as well as the fastest perceptual estimates are biased towards the vector average of the edge motions. It is only after a few hundred milliseconds of exponential direction error decay that the eyes correctly track the object or that human subjects report the correct direction of motion.

3.2 Gratings

Gratings Sizes. In [4] the authors use a drifting grating viewed through a circular aperture. The orientation of the grating is constant and orthogonal to its drifting direction, but the diameter of the circular aperture varies among the stimuli. The authors quantify the change in eye direction during several time windows with respect to the diameter of the aperture. Their goal is to provide a quantitative measure of the spatial summation area, i.e. the smallest diameter leading to the strongest change in eye position. Such spatial summation functions can be seen as a global readout of the motion integration performed in area MT. It is however also possible to look at the perceptual effects of such stimuli: varying sizes of grating patches affect motion detection as well as motion after effect. Many psychophysical studies have been conducted on the perceptual consequences of the centre-surround interactions in early visual areas (see [32] for a review) and it becomes possible to compare these results for the properties of neuronal receptive fields in area V1, MT, or MST in macaque monkeys.

Barber Pole. In the classical barber pole illusion, a translating grating is viewed through a rectangular aperture, leading to two orthogonal sets of 2D

cues [38]. The larger set of 2D cues originates from the longest side of the rectangular aperture, while the smaller set of 2D cues originates from the shortest side. According to psychophysical experiments, as well as neurobiological data, the final perceived motion direction is the same as the orientation of the elongated side of the aperture, after an initial direction orthogonal to the grating orientation [23]. The perceived motion direction thus corresponds to the 2D cues with the greater number of occurrences.

Again, similar observations are available at psychophysical [12, 19] and neuronal [27] levels. It is thus possible to compare model output with a global readout such as time-dependant ocular pursuit but also to compare the dynamics of single model neurons with that of V1 and MT neurons.

4 Results

4.1 Readouts Definition

Common output is necessary in order to compare models together and abstract all implementation dependent issues coming from the large variety of motion models. For example, if eye-movement like output can be defined, then models can be compared together in term of dynamics. As illustrated by this example, our goal here is to propose output formats corresponding to classical readouts as defined in psychophysics. Our goal is to describe for each readout, what they are supposed to measure, and how they are measured in psychophysics.

Since the notion of local measurement has no clear interpretation in term of neural architecture or activity, the readouts defined herein correspond to global motion estimations. For example, it is known that the preferred motion of neurons in both V1 and MT changes depending on the stimulus [28], or that the perceived motion and the neural activity can differ [17]. Considering not only the local estimates but also more distant features makes an important difference with classical computer vision methodology for optical flow where only precision of local estimates matters.

Our goal is to define qualitatively which common outputs are needed from models (i.e. readout inspired from real neuroscience experiments). However no general formula for readouts can be given for three main reasons. The first reason is that the cortical mechanisms leading to a readout from neural activity are usually not clearly established and *a fortiori* it is hard to model them rigorously. The second reason is that the variety of motion representations in models makes it impossible to write a general formula that would be valid for any kind of representation. The last reason is that readouts defined here are sometimes inherent to models. Indeed some models already provide an eye-movement output [25] whereas others considers neural activity in cortical areas [6].

In this article two kinds of evaluation are considered: The *static* evaluation considers only the result at convergence whereas the *dynamic* evaluation focuses on the dynamics of motion integration.

For the *dynamic evaluation*, the readout is expressed as a time independent value, such as a perceived motion direction. Given a stimulus, some experiments

require subjects to give their perception concerning the motion they perceive. In general, this perceived motion readout ignores time evolution. From a modelling point of view, the perceived motion readout can be a global velocity corresponding to the steady state. For the 2D motion integration tasks, we can assume that perceived motion corresponds to the final output from eye movement readout.

For the *static evaluation*, the readout is expressed as a time dependant value, such as smooth pursuit eye movements. Voluntary eye movement to track motion are directly related to our interpretation of the scene in term of motion content. Primates use two types of voluntary eye movement to track objects of interest: smooth pursuit and saccades. Pursuit eye movements are driven by visual motion and rely on both low-level and high-level motion processing. Pursuit initiation is critically dependent upon visual motion processing in cortical areas MT and MST. It presents the interest of being a simple motor responses that requires an accurate estimate of the global direction and speed of a single object, despite its properties such as shape or colour. It is therefore a good probe of object motion processing and in particular it reflects many of the dynamical properties of low level motion computation. From a modelling point of view, smooth pursuit eye movement is a single time-dependent vector, and we only consider the eye direction since speed is generally ignored in experiments.

4.2 Scoring Procedure

The full scoring procedure for each class of stimuli is available online. For each stimulus, instructions are detailed (see for example Figure 2 for translating diamonds stimuli). In order to show the general idea, let us explain what is the scoring procedure for static evaluation of the translating diamonds stimuli. Our reference paper for this case will be [24] as it presents ocular following measurements that we can use in our evaluation.

For a given an approach, our evaluation procedure starts from the estimated global motion direction at every frame and for each of the stimuli in this class, i.e. for translating diamonds translating in one out of the four possible directions (right, up, left, down) and oriented either clockwise (CW) or counter clockwise (CCW) . Let us denote by $e_{\mathcal{S}}(t)$ the estimated global direction dynamics for a stimulus $\mathcal{S}$, with $t \in [0, 450\,\text{ms}]$ and $\mathcal{S} \in \{\text{right,up,left,down}\} \times \{\text{CW,CCW}\}$. From this global estimated direction, the instantaneous direction error $\varepsilon_{\mathcal{S}}(t)$ is defined by

$$\varepsilon_{\mathcal{S}}(t) = e_{\mathcal{S}}(t) - \hat{e}_{\mathcal{S}}(t),$$

where $\hat{e}_{\mathcal{S}}(t)$ is the true object motion. Some results are shown in Figure 3 (a) for the biologically inspired artificial vision model proposed in [35]. Here the estimated global motion direction was obtained from the MT layer activity.

Our goal is to compare this estimated direction error $\varepsilon_{\mathcal{S}}(t)$ to the direction error observed with human subjects (see Figure 3 (b) from [24]). We followed the same procedure as the one defined in [24]. Estimated direction error is fitted with the function.

Translating diamonds: static

Masson, G.S. and Stone, L.S.
J Neurophysiol, 2002.

1. For each stimulus, compute global direction

a. The static (final) global motion direction is computed for each stimulus

$d_s \in [0, 2\pi[\qquad \text{with } s \in \{\text{up}, \text{down}, \text{left}, \text{right}, \} \times \{\text{CW}, \text{CCW}\}$

2. For each stimulus, we compute a direction error

a. We know the true object motion direction e_s for each stimulus.

b. We can compute the difference between the two

$$e_s = \hat{d}_s - d_s$$

3. We score the direction errors

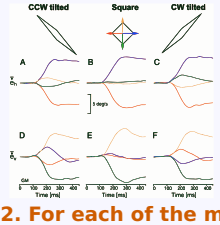
a. Assuming a normal law with data from the literature, where $\hat{e}_s$ is the average direction error, and σ_s its deviation:

$$k_s = G_{\sigma_s}(e_s - \hat{e}_s)$$

Translating diamonds: dynamic

Masson, G.S. and Stone, L.S.
J Neurophysiol, 2002.

1. For each stimulus, extract eye direction error



a. The eye direction dynamics is denoted for each stimulus s by:

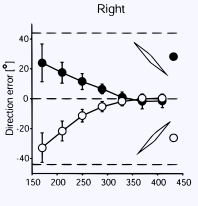
$$e_s(t) \in [0, 2\pi)$$

$$\text{with } t \in [0, 450],\ s \in \{\text{up}, \text{down}, \text{left}, \text{right}, \} \times \{\text{CW}, \text{CCW}\}$$

b. From the veridical object motion, $\hat{e}_s$, we compute a direction error ϵ_s:

$$\epsilon_s(t) = \hat{e}_s(t) - e_s(t)$$

2. For each of the motion direction, we fit the error



a. For a given motion direction, we average absolute errors for the clockwise and counter clockwise orientations.

$$h_d(t) = 0.5\,\text{abs}\left(e_{d,\text{CW}}(t) - e_{d,\text{CCW}}(t)\right)$$

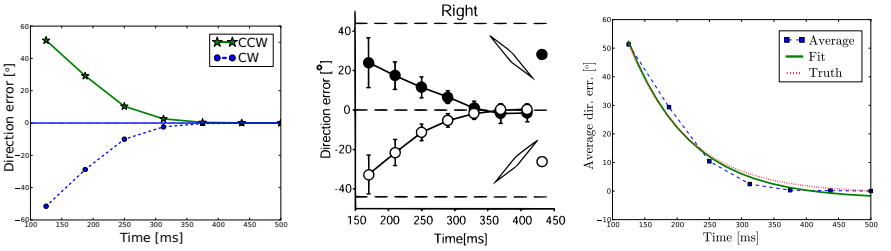
b. We fit the result with the following function

$$f_{A,B,\tau}(t) = A\,\exp\left(-\tfrac{t}{\tau}\right) + B$$

3. We score the fitted parameters

Assuming a normal law with data from the literature.

Fig. 2. Example of slides describing the scoring procedure (for the translating diamonds stimuli) which are available on the benchmark website



(a) (b) (c)

Fig. 3. Scoring procedure for translating diamonds (rightward-motion) (a) Estimated direction error of the biologically inspired artificial vision model proposed in [35]. (b) Oculo-motor dynamics recorded from human subjects. Reproduced from [24]. (c) Comparison between the estimated direction error for all stimuli from the class (Average), the exponential fit on this estimated direction error (Fit) and the fit on the experimental data (Truth).

$$f_{\alpha,\beta,\tau}(t) = \alpha \, \exp\left(-t/\tau\right) + \beta,$$

where α, β, and τ are the fitting parameters to be adjusted. These parameters can then be compared to what is obtained with human observers. In [24] the authors estimated the average values and standard deviations for these parameters (denoted respectively by $\mu(\hat{\eta})$ and $\sigma(\hat{\eta})$ for a parameter $\eta \in \{\alpha, \beta, \tau\}$). Assuming a Gaussian distribution for parameters coming from human subjects and given a stimulus $\mathcal{S}$, we defined a score $s \in [0, 1]$ for each parameter by:

$$s_\eta = \exp(-(\eta - \mu(\hat{\eta}))^2/\sigma(\hat{\eta})^2), \qquad \text{with } \eta \in \{\alpha, \beta, \tau\}.$$

Finally, a global score can be obtained by averaging scores over all stimuli and parameters.

4.3 Baseline Results

We applied our evaluation methodology to both biologically inspired artificial vision models [6, 35] and computer vision models [16, 21, 34, 11] by running either the original implementation from the authors or the code that was available in the OpenCV library [10]. A single set of parameters were experimentally tuned in order to achieve the overall best score across all experiments. As defined in section 4.1, we discuss below the results obtained in the static and dynamic evaluations. For each scenario, results are presented into tables by scores between zero (low performance) and one (high performance). Algorithms are ranked according to their average score across all experiments. Complete details for each stimuli and evaluation procedure can be found on the associated website.

Static Evaluation. Results are presented in Table 1. As a general comment, it is interesting to remark that models performance somewhat follows research

Table 1. Static evaluation results. For each approach and each experiment a score between 0 (worse) and 1 (best) is given depending on the final motion direction error. BMOCV denotes the block matching algorithm found in the OpenCV library.

Approach	Avg.				
TMK·10 [35]	1.00	1.00	1.00	1.00	1.00
SRDB·10 [34]	0.86	1.00	1.00	0.65	0.78
BM·10 [11]	0.74	1.00	1.00	0.00	0.98
BN·07 [6]	0.68	1.00	1.00	0.36	0.38
LK·81 [21]	0.45	0.81	0.00	0.99	0.00
HS·81 [16]	0.39	0.52	0.00	1.00	0.03
BMOCV [10]	0.19	0.00	0.32	0.44	0.00

Table 2. Dynamic evaluation results. For each approach and each experiment a score between 0 (worse) and 1 (best) is given depending on the fitting procedure described in the main text.

Approach	Avg.			
TMK·10 [35]	0.68	1.00	0.96	0.08
LK·81 [21]	0.37	0.75	0.36	0.00
SRDB·10 [34]	0.37	0.75	0.36	0.00
BM·10 [11]	0.35	0.50	0.36	0.18
BN·07 [6]	0.32	0.50	0.39	0.07
BMOVC [10]	0.31	0.50	0.36	0.05
HS·81 [16]	0.26	0.75	0.03	0.00

evolution. For example the seminal approaches for optical flow proposed by Horn and Schunck [16] or Lucas and Kanade [21] show quite a poor performance on most stimuli. The fact that these approaches are differential and not multi-scale largely explains this performance. Being differential, the optical flow is estimated based on the brightness consistency assumption, which is a local indication. Thus when there is a majority of 1D cues, with hardly no texture, the input to differential algorithms is not very informative and leads to an aperture problem that is hard to solve numerically.

Considering multi-scale approaches is today one classical method to solve the aperture problem more efficiently. This solution is now used by most current models, such as the recent models by [34, 11], which are now among the best computer vision models (see the latest results online from [2]). Interestingly, those models also perform well for most of our experiments.

Biologically inspired artificial vision models [6, 35] show high performance. In particular the model proposed in [35] obtains the maximum score. The major strength of this model is that its design is naturally multi-scale as it is inspired from the multi-layer architecture of the brain cortical areas (V1 and MT) with proper connectivity patterns. This is one important result of this evaluation methodology because it allows to show that taking biology into account can lead to extra performance.

Dynamic Evaluation. Results are presented in Table 2. Studying the dynamical properties of motion integration is quite a new topic and very few biological data (psychophysical, oculo-motor, or neural dynamics) were available when most models were proposed. For this reason, most models were often static, i.e. they were interested in estimating the final percept or optical flow, ignoring how the solution evolves in time to the final percept.

Studying the dynamics has mainly been considered in [35]. This model shows the best performance on translating bars and diamonds, considering the recent neuronal, psychophysical, and behavioural findings [26, 27, 24, 37]. For all other models, since there is no true dynamics, scores are not very informative. Yet, dynamics remain an open issue in the modelling community. None of models tested here performed correctly on the gratings stimulus, suggesting that the underlying mechanisms remain to be found.

5 Conclusion

In this article we set the basis for a new evaluation methodology for motion models which is based on human performance. This work generalises in a rigorous way the evaluation procedures done for most motion models proposed in computational neuroscience. By carefully defining a unified database and proper scoring procedures, it is now possible to perform non-biased comparisons between models, since stimuli are not optimised for a given approach. Our database is freely available on the web together with scoring instructions and results. We provided baseline results for both biologically inspired artificial vision models and state-of-the-art computer vision models. Our results also contribute to show what is the interest to consider biologically inspired models in the computer vision community. By considering stimuli from the psychophysics community, one can further challenge motion models, in addition to existing optical benchmarks.

The proposed evaluation methodology can of course be extended. In this study several properties affecting the motion integration mechanisms were ignored. For instance, disparity used in binocular experiments is missing, and thus it could be probably possible to evaluate other kinds of models where motion and depth are combined [7, 39, 30, 33]. Another property ignored herein is the contrast. In a wide range of psychophysical and neurobiological stimuli, contrast has a considerable effect on motion integration. For instance the receptive field size in areas V1 MT changes with contrast [31, 29]. Contrast also influences behavioural results [24, 37, 9]. Those kinds of stimuli variations should also be considered in

a more comprehensive data set. One could also think about other stimuli. For instance, it is well known that most of the motion stimuli are multi-stable. In the case of drifting plaids one can perceive either two gratings with different velocities, or one single plaid motion [18]. Incorporating this multi-stability in models is still only at the sketch level in models [14, 36, 35], and mostly ignored in motion benchmarks.

Acknowledgements. The authors particularly thank Pascal Mamassian for enriching thoughts related to this work. This research work has received funding from the Région PACA, the CNRS, the European Community (FACETS, IST-FET, Sixth Framework, N°025213, and SEARISE, Seventh Framework, N°215866), and the Agence Nationale de la Recherche (ANR, NATSTATS).

References

[1] Adelson, E.H., Bergen, J.R.: Spatiotemporal energy models for the perception of motion. Journal of the Optical Society of America A 2(2), 284–299 (1985)
[2] Baker, S., Scharstein, D., Lewis, J.P., Roth, S., Black, M.J., Szeliski, R.: A database and evaluation methodology for optical flow. In: International Conference on Computer Vision, ICCV 2007 (2007)
[3] Barron, J.L., Fleet, D.J., Beauchemin, S.S.: Performance of optical flow techniques. IJCV 12(1), 43–77 (1994)
[4] Barthelemy, F.V., Vanzetta, I., Masson, G.S.: Behavioral receptive field for ocular following in humans: dynamics of spatial summation and center-surround interactions. Journal of Neurophysiology 95(6), 3712 (2006)
[5] Bayerl, P., Neumann, H.: Disambiguating visual motion through contextual feedback modulation. Neural Computation 16(10), 2041–2066 (2004)
[6] Bayerl, P., Neumann, H.: Disambiguating visual motion by form-motion interaction–A computational model. International Journal of Computer Vision 72(1), 27–45 (2007)
[7] Beck, C., Neumann, H.: Interactions of motion and form in visual cortex – a neural model. Journal of Physiology - Paris 104, 61–70 (2010)
[8] Biber, U., Ilg, U.J.: Initiation of smooth-pursuit eye movements by real and illusory contours. Vision Research 48(8), 1002–1013 (2008)
[9] Born, R.T., Pack, C.C., Ponce, C., Yi, S.: Temporal evolution of 2-dimensional direction signals used to guide eye movements. Journal of Neurophysiology 95, 284–300 (2006)
[10] Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000)
[11] Brox, T., Malik, J.: Large displacement optical flow: descriptor matching in variational motion estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence (2010)
[12] Castet, E., Charton, V., Dufour, A.: The extrinsic/intrinsic classification of two-dimensional motion signals with barber-pole stimuli. Vision Research 39(5), 915–932 (1999)
[13] Cesmeli, E., Lindsey, D.T., Wang, D.L.: An oscillatory correlation model of human motion perception. In: Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, IJCNN 2000, vol. 4 (2000)
[14] Giese, M.: Dynamic Neural Field Theory for Motion Perception. Springer (1998)

[15] Grossberg, S., Mingolla, E., Pack, C.: A neural model of motion processing and visual navigation by cortical area MST. Cerebral Cortex 9(8), 878–895 (1999)
[16] Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artificial Intelligence 17, 185–203 (1981)
[17] Huang, X., Albright, T.D., Stoner, G.R.: Stimulus dependency and mechanisms of surround modulation in cortical area MT. Journal of Neuroscience 28(51), 13889 (2008)
[18] Hupé, J.M., Rubin, N.: The dynamics of bi-stable alternation in ambiguous motion displays: a fresh look at plaids. Vision Research 43(5), 531–548 (2003)
[19] Kooi, T.L.: Local direction of edge motion causes and abolishes the barberpole illusion. Vision Research 33(16), 2347–2351 (1993)
[20] Lorenceau, J., Shiffrar, M., Wells, N., Castet, E.: Different motion sensitive units are involved in recovering the direction of moving lines. Vision Research 33(9), 1207–1217 (1993)
[21] Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: Proceedings of Imaging Understanding Workshop, pp. 121–130 (1981)
[22] Masson, G.S., Ilg, U.J. (eds.): Dynamics of Visual Motion Processing Neuronal, Behavioral, and Computational Approaches, 1st edn. Springer (2010)
[23] Masson, G.S., Rybarczyk, Y., Castet, E., Mestre, D.R.: Temporal dynamics of motion integration for the initiation of tracking eye movements at ultra-short latencies. Visual Neuroscience 17(05), 753–767 (2000)
[24] Masson, G.S., Stone, L.S.: From following edges to pursuing objects. Journal of Neurophysiology 88(5), 2869 (2002)
[25] Montagnini, A., Mamassian, P., Perrinet, L., Castet, E., Masson, G.S.: Bayesian modeling of dynamic motion integration. Journal of Physiology-Paris 101(1-3), 64–77 (2007)
[26] Pack, C.C., Born, R.T.: Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. Nature 409, 1040–1042 (2001)
[27] Pack, C.C., Gartland, A.J., Born, R.T.: Integration of contour and terminator signals in visual area MT of alert macaque. The Journal of Neuroscience 24(13), 3268–3280 (2004)
[28] Pack, C.C., Hunter, J.N., Born, R.T.: Contrast dependence of suppressive influences in cortical area MT of alert macaque. Journal of Neurophysiology 93(3), 1809 (2005)
[29] Pack, C.C., Hunter, J.N., Born, R.T.: Contrast dependence of suppressive influences in cortical area MT of alert macaque. Journal of Neurophysiology 93(3), 1809–1815 (2005)
[30] Salgado, A., Sánchez, J.: Temporal Constraints in Large Optical Flow Estimation. In: Moreno Díaz, R., Pichler, F., Quesada Arencibia, A. (eds.) EUROCAST 2007. LNCS, vol. 4739, pp. 709–716. Springer, Heidelberg (2007)
[31] Sceniak, M.P., Ringach, D.L., Hawken, M.J., Shapley, R.: Contrast's effect on spatial summation by macaque V1 neurons. Nature Neuroscience 2(8), 733–739 (1999)
[32] Seriès, P., Georges, S., Lorenceau, J., Frégnac, Y.: A network view of the structure of center/surround modulations of V1 receptive field properties in visual and cortical spaces. Neurocomputing 38, 881–888 (2001)
[33] Strecha, C., Van Gool, L.: Motion - Stereo Integration for Depth Estimation. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part II. LNCS, vol. 2351, pp. 170–185. Springer, Heidelberg (2002)

[34] Sun, D., Roth, S., Darmstadt, T.U., Black, M.J.: Secrets of Optical Flow Estimation and Their Principles. In: IEEE Int. Conf. on Computer Vision and Pattern Recognition (2010)
[35] Tlapale, É., Masson, G.S., Kornprobst, P.: Modelling the dynamics of motion integration with a new luminance-gated diffusion mechanism. Vision Research 50(17), 1676–1692 (2010)
[36] Veltz, R., Faugeras, O.: Local/global analysis of the stationary solutions of some neural field equations. SIAM J. Applied Dynamical Systems 9(3), 954–998 (2010)
[37] Wallace, J.M., Stone, L.S., Masson, G.S.: Object motion computation for the initiation of smooth pursuit eye movements in humans. Journal of Neurophysiology 93(4), 2279–2293 (2005)
[38] Wallach, H.: Über visuell wahrgenommene Bewegungsrichtung. Psychological Research 20(1), 325–380 (1935)
[39] Werlberger, M., Trobin, W., Pock, T., Wedel, A., Cremers, D., Bischof, H.: Anisotropic Huber-L1 optical flow. In: Proceedings of the British Machine Vision Conference, BMVC (2009)