ProbFlow: Joint Optical Flow and Uncertainty Estimation

Anne S. Wannenwetsch\textsuperscript{1} 
Margret Keuper\textsuperscript{2} 
Stefan Roth\textsuperscript{1}

\textsuperscript{1}TU Darmstadt 
\textsuperscript{2}University of Mannheim

Abstract

Optical flow estimation remains challenging due to untextured areas, motion boundaries, occlusions, and more. Thus, the estimated flow is not equally reliable across the image. To that end, post-hoc confidence measures have been introduced to assess the per-pixel reliability of the flow. We overcome the artificial separation of optical flow and confidence estimation by introducing a method that jointly predicts optical flow and its underlying uncertainty. Starting from common energy-based formulations, we rely on the corresponding posterior distribution of the flow given the images. We derive a variational inference scheme based on mean field, which incorporates best practices from energy minimization. An uncertainty measure is obtained along the flow at every pixel as the (marginal) entropy of the variational distribution. We demonstrate the flexibility of our probabilistic approach by applying it to two different energies and on two benchmarks. We not only obtain flow results that are competitive with the underlying energy minimization approach, but also a reliable uncertainty measure that significantly outperforms existing post-hoc approaches.

1. Introduction

Optical flow estimation has been extensively studied for more than three decades [5, 9, 14, 21, 36]. However, motion boundaries, large displacements, and occlusions still lead to erroneous flow fields, especially in challenging imaging conditions. In contrast, flow predictions are almost errorless in textured regions of uniform motion, which causes the reliability of optical flow predictions to vary greatly across the image. Uncertainty measures\textsuperscript{1} aim to predict the reliability of the flow estimates and rate each estimate according to its predicted accuracy [24, 27, 30]. One application of such measures is to improve optical flow estimation using different ways of post-processing [4, 36]. Moreover, uncertainty measures represent an important tool when optical flow is used as a cue for other computer vision tasks, such as image segmentation or tracking, which consequently depend on the correctness of the flow. In such cases, error propagation can be avoided if only flow estimates of high confidence are taken into consideration, e.g. [33, 47].

Several post-hoc confidence measures have been introduced to directly estimate the uncertainty of flow predictions, e.g. [3, 20, 24, 30]. However, optical flow estimation is known to be rather challenging as a one-shot process [14, 15, 43]. We thus believe it is natural to ask why confidence estimation should be restricted to non-iterative (post-hoc) approaches. Moreover, flow algorithms for a concrete problem are carefully selected for the task at hand and provide us with an underlying model. We, thus, argue that it is very desirable to apply a model-inherent uncertainty measure tailored to the chosen method. We even go one step further and aim to jointly predict optical flow and a corresponding uncertainty in order to preserve and extract important information contained in the flow estimation process.

Our work focuses on energy-based methods [3, 15, 43] as they represent a flexible framework in widespread use. Starting from a general energy formulation, we derive a probability distribution of the flow field conditioned on the input images. In this setting, optical flow can be determined by minimizing the expected loss under the modeled posterior and the corresponding uncertainty measure is naturally obtained as the marginal entropy of the posterior at every pixel. Due to the intractability of exact posterior inference in general, we rely on variational approximate inference and

---

\textsuperscript{1} Uncertainty measures are closely related to confidence measures as their values are inversely related to confidence estimates.
use a mean-field approach. In this way, flow predictions and uncertainty estimations can be obtained as the result of a joint minimization problem that makes an additional training step of the uncertainty measure unnecessary [17, 30]. Fig. 1 shows an exemplary flow field and uncertainty prediction of our probabilistic approach obtained on the Sintel benchmark [11]. We term our approach ProbFlow.

To the best of our knowledge, we introduce the first fully probabilistic flow estimation approach that shows competitive results on established benchmarks. Moreover, our work is unique in that it is broadly applicable to a large number of existing energy-based optical flow algorithms and that we demonstrate how to benefit from the probabilistic framework in the context of uncertainty measures.

We apply ProbFlow to a classic Horn-Schunck-style energy as proposed in [36] and to the more recent EpicFlow formulation [36]. We show that the obtained flow fields are competitive with or even outperform the corresponding energy minimization results on the Middlebury [2] and Sintel [11] datasets. To assess the performance of the uncertainties estimated by our method, we rely on existing evaluation approaches and propose a new criterion based on the Spearman’s rank correlation coefficient. Our uncertainty measure is clearly superior in comparison to various competing approaches and significantly outperforms the best existing methods. We further show the benefits of our uncertainty estimates in an application to motion segmentation [22] by generating highly reliable point trajectories.

2. Related Work

Starting from the work of Horn and Schunck [21], global energy minimization approaches have been used extensively for optical flow estimation; see [3, 15, 43] for indepth overviews of existing methods and best practices. Despite their conceptual simplicity, global energy formulations are still up-to-date [13, 29]. Moreover, energy-based approaches are frequently used for post-processing optical flow predictions obtained through other means [1, 14, 36].

In the past, a variety of optical flow uncertainty measures have been proposed. Various simple methods are based on the characteristics of the input data only, e.g., [3, 20]. We refer to [24, 27] for a more comprehensive summary of such confidence approaches. In comparison to our work, all of these measures omit important sources of information by not considering the estimated optical flow field itself.

A second class of confidence measures relies on an analysis of the estimated flow fields. Kondermann et al. [23] learn a linear subspace of true spatio-temporal flow neighborhoods and use the reconstruction error of an estimated flow vector to evaluate its reliability. In [24], a probabilistic model of flow patches is learned and approximated with a Gaussian distribution, which yields a confidence measure based on hypothesis testing. But this second class of confidence measures does not consider all aspects of flow uncertainty either, as the input images are not taken into account.

Recently, Mac Aodha et al. [30] proposed a confidence measure based on the input as well as the output of an optical flow algorithm. Using a multi-cue feature vector, a classifier is trained to predict whether the endpoint error at a certain pixel is smaller than a previously defined threshold. The probability output of the classifier is used as a confidence measure. However, [30] does not take advantage of uncertainty information available in the flow estimation procedure itself and requires a separate training step that is unnecessary for our model-inherent approach.

Uncertainty measures tailored to energy minimization approaches have been proposed especially for local energy methods, e.g., [3, 31, 40, 41]. Model-inherent confidence measures designed for global energy formulations are quite rare, on the other hand. Kybic and Nieuwenhuis [27] introduce a method that relies on bootstrap resampling. Based on varying pixel contributions, repeated optical flow estimation is performed and a confidence measure is determined as the total standard deviation of the obtained flow predictions. Bruhn and Weickert [8] propose the inverse local energy as a confidence measure, which is applicable to a broad variety of energies. A large uncertainty is thus associated with a strong violation of the local model assumptions encoded in the energy formulation. Gehrig and Scharwächter [17] use the energy and combine it with several features such as the spatial and temporal flow variance in order to obtain a real-time confidence estimate. In contrast to the two above approaches, our method does not explicitly limit the receptive field of the confidence measure by only considering the local neighborhood of a pixel. Instead, uncertainty propagation is facilitated by means of iterative spatial inference.

In the past, several methods based on a probabilistic flow formulation have been proposed [35, 39, 40, 41]. However, many of these works are based on simplifying assumptions such as locally constant flow or a Gaussian distribution of brightness constancy errors. In [19, 26, 44], probabilistic approaches are applied to obtain improved models of optical flow. In contrast to our work, these methods do not apply a fully probabilistic approach, but fall back to a maximum a-posteriori (MAP) estimate, i.e. minimize the underlying energy. Glocker et al. [18] use flow uncertainties in a dynamic Markov random field but rely on a discrete approach. The most closely related work of Chantas et al. [12] applies a variational-Bayes approach similar to ours. However, their method is not designed as a stand-alone flow algorithm, but only as an improved initialization in comparison to a simple Horn-Schunck approach. Therefore, the obtained results are not competitive with respect to the state of the art. Moreover, the paper does not take advantage of the probabilistic approach in the context of uncertainty measures.
3. Energy Framework

Since our joint, model-inherent uncertainty estimation developed below is broadly applicable to different kinds of energy-based optical flow approaches [3, 15, 43], we first introduce a formalization that allows us to derive previously proposed optical flow methods in a unified manner.

In the following, we estimate optical flow between an image pair \( I = \{I_1, I_2\} \) and denote the estimate as \( y = (y_{ij})_{i,j} = (u_{ij}, v_{ij})^T \) for pixels \((i,j), \ i = 1, \ldots, n, j = 1, \ldots, m\). Energy-based approaches estimate the optical flow as the minimizer \( y^* \) of an energy function

\[
E(y; I) = E_D(y; I) + \lambda_S E_S(y)
\]

with \( E_D(y; I) \) denoting a data term that encourages the flow to be consistent with input images \( I_1 \) and \( I_2 \). The spatial term \( E_S(y) \) imposes a (smoothness) prior on the flow, and \( \lambda_S \) represents a trade-off parameter between the terms.

In the following, we use functions \( f_D(\cdot) \) and \( f_S(\cdot) \) to formalize violations of the assumptions underlying the chosen optical flow model. For instance, the intensity difference can then be described as a sum of the contributions from all pixels such that

\[
E_D(y; I) = \sum_{i,j} \rho_D(f_D(y_{ij}; I))
\]

\[
E_S(y) = \sum_{i,j} \sum_{(i',j') \in S(i,j)} \rho_S(f_S(y_{ij}, y_{i'j'}))
\]

with \( S(i,j) \) describing a set of neighbors of pixel \((i,j)\).

It has been shown that the estimated flow field \( y^* \) can be improved by adding a non-local term to the energy function in Eq. (1) [25, 43]. Thus, we optionally use an additional non-local term \( E_N(y) \) akin to \( E_S(y) \), considering an extended neighborhood \( N(i,j) \). Again, \( E_N(y) \) is described by a model assumption \( f_N(\cdot) \) and its corresponding penalty function \( \rho_N(\cdot) \).

Common choices for penalty functions \( \rho(\cdot) \) are, e.g., a quadratic function [21], a Lorentzian function [5], or a (generalized) Charbonnier function [9, 43]. Many of these functions can be described by the negative logarithm of a Gaussian Scale Mixture (GSM) [44, 46]. Thus, in the following we rely on this simple but powerful class of functions and represent the penalty terms as GSMs of \( L \) components

\[
\rho(z) = -\log \left[ \sum_{l=1}^{L} \pi_l \mathcal{N}(z; 0, \sigma_l^2) \right]
\]

with \( \mathcal{N}(z; \mu, \sigma) \) being a normal distribution with mean \( \mu = 0 \) and variance \( \sigma = \sigma_l \); the weights \( \pi_l \geq 0 \) sum to 1. As we will see, GSMs also benefit our probabilistic approach.

4. Probabilistic Interpretation and Inference

We now aim to approach optical flow estimation in a probabilistic manner. As the energy function \( E(y; I) \) from Eq. (1) describes a Markov random field, it is easy to derive the corresponding posterior distribution in its Gibbs form as

\[
P(y \mid I) = \frac{1}{Z} \exp \left\{ -\frac{1}{T} E(y; I) \right\}
\]

with partition function \( Z = Z(I, T, \lambda_S, \lambda_N) \) and temperature \( T \). In the following, w.l.o.g. we set \( T = 1 \) and introduce an additional parameter \( \lambda_D \) scaling the data term \( E_D \).

To ease probabilistic inference, we use the same procedure as [43] and introduce an auxiliary flow field \( \hat{y} \) that allows us to decouple the non-local potential from the remaining terms of the posterior distribution in Eq. (5), i.e.,

\[
p(y, \hat{y} \mid I) = \frac{1}{Z} \exp \left\{ -\lambda_D E_D(y; I) - \lambda_S E_S(y) \right\}
\]

with \( E_C(y, \hat{y}) = \sum_{i,j} \|y_{ij} - \hat{y}_{ij}\|^2 \).

The log-posterior of Eq. (6) now has terms in the form of the logarithm of a sum of exponentials, which is challenging when deriving closed-form mean-field updates below. Here, we benefit from our choice to model each penalty function \( \rho(\cdot) \) as a GSM. We follow [16, 28] to retain explicit latent variables \( h = \{h_D, h_S, h_N\} \), which are chosen to follow a discrete distribution and to have a 1-of-L representation. We then obtain an augmented penalty function \( \rho_\gamma(z, h_\gamma) \) with \( \gamma \in \{D, S, N\} \) as

\[
\rho_\gamma(z, h_\gamma) = -\log \left[ \prod_{l=1}^{L} \pi_l^{h_{\gamma,l}} \mathcal{N}(z; 0, \sigma_l^{h_{\gamma,l}}) \right].
\]

At this point, the posterior \( P(y, \hat{y}, h \mid I) \) represents the probabilistic equivalent of the energy \( E(y; I) \) in Eq. (1).

In our probabilistic setup, the flow estimate \( y^* \) is chosen to minimize the expected loss over \( p(y, \hat{y}, h \mid I) \), i.e.,

\[
y^* = \arg \min_{\hat{y}} \mathbb{E}_{p(y, \hat{y}, h \mid I)}[l(y, \hat{y})]
\]

with \( l(\cdot, \cdot) \) being a suitable loss function such as the Average End-Point Error (AEPE) [34]. The desired uncertainty measure can be obtained by considering the marginal distribution \( p(y_{ij} \mid I) \) of the flow estimates. In particular, we propose to use the marginal entropy at every pixel as a model-inherent uncertainty estimate of the flow prediction.

Inference. To obtain a flow estimate and the underlying marginal distributions at every pixel, we rely on an approximate inference scheme. We use variational inference [45].

\[\text{Not to be confused with variational formulations common in flow.}\]
and approximate \( p(y, \hat{y}, h \mid I) \) with the help of a tractable distribution \( q(y, \hat{y}, h ; \theta) \) and variational parameters \( \theta \).

Following a naive mean-field assumption, we choose the parametric distribution \( q \) to be factorized over \( y, \hat{y}, h \), as well as \( h \). The marginal distribution of flow vectors \( y_{ij} \) and \( \hat{y}_{ij} \) is assumed to be Gaussian, e.g.

\[
q(y_{ij} ; \theta) = \mathcal{N}(y_{ij} ; \mu_{ij}, \Sigma_{ij}) .
\]

It is reasonable to assume the horizontal and vertical flow components to be uncorrelated [37]. Thus, the covariances \( \Sigma_{ij} \) are modeled as diagonal matrices. As in [16, 28], the distributions of the latent variables are chosen to be multinomial

\[
q(h_{\gamma,ij} ; \theta) = \prod_{l=1}^{L} k_{\gamma,ij,l} \]

so that the parameters \( k_{\gamma,ij,l} \geq 0 \) satisfy \( \sum_{l} k_{\gamma,ij,l} = 1 \). The approximating distribution \( q \) is then given as

\[
q(y, \hat{y}, h ; \theta) = \prod_{i,j} \left[ q(y_{ij} ; \theta) \cdot q(\hat{y}_{ij} ; \theta) \prod_{\gamma \in \{D,S,N\}} q(h_{\gamma,ij} ; \theta) \right] \tag{11}
\]

with \( \theta = \{ \mu_{ij}, \nu_{ij}, \Sigma_{ij}, \Sigma_{ij,k}, \kappa_{ij,\gamma} \} \). Suitable variational parameters \( \theta^* \) of \( q \) are determined such that the Kullback-Leibler (KL) divergence between \( p \) and its approximating distribution \( q \) is minimized, i.e.

\[
\theta^* = \arg \min_{\theta} D_{KL}(q(y, \hat{y}, h ; \theta) \mid p(y, \hat{y}, h \mid I)) \tag{12}
\]

\[
= \arg \min_{\theta} E_{q(y, \hat{y}, h ; \theta)} \left[ \log q(y, \hat{y}, h ; \theta) \right] - E_{q(y, \hat{y}, h ; \theta)} \left[ \log p(y, \hat{y}, h \mid I) \right]. \tag{13}
\]

Due to the usage of explicit latent variables \( h \), it is possible to compute the expectations in Eq. (13) in an analytic way. While the derivation of the corresponding equations is tedious, the individual steps are elementary; see supplemental material for a more detailed explanation of the procedure.

We now estimate the flow by replacing the posterior \( p \) in Eq. (8) with its approximating distribution \( q \). When performing Bayesian risk minimization of the AEPE, the optical flow prediction is obtained as \( y_{ij}^* = \mu_{ij} \) and, therefore, corresponds to the mode of the variational distribution \( q \).

As \( q(y, \hat{y}, h ; \theta) \) was defined in a factorized way, the corresponding marginal distribution at pixel \( (i,j) \) is given as \( q(\hat{y}_{ij} ; \theta) \) and our proposed model-inherent uncertainty measure can be obtained as the marginal entropy

\[
\Psi_{\text{ProbFlow}} = H(y_{ij}) = \log \left( \det (\Sigma_{ij}) \right) + \text{const.} \tag{14}
\]

5. Specific Models

We now apply our ProbFlow approach to two specific energy functions commonly used for optical flow estimation.

5.1. Probabilistic Classic Flow

We first consider a classical Horn-Schunck-based objective based on brightness constancy as given in [43]. Following the common approach to use a first-order Taylor approximation for a linearization of the data term, we obtain

\[
f_{\text{D}}(y_{ij} ; I) = I_{2} (i + u_{ij}^{0}, j + v_{ij}^{0}) - I_{1} (i,j) + \nabla_{2} I_{2} (i + u_{ij}^{0}, j + v_{ij}^{0})^T (y_{ij} - y_{ij}^{0}) \tag{15}
\]

where \( y_{ij}^{0} \) is the point of approximation and \( \nabla_{2} I_{2} \) denotes the spatial derivatives of \( I_{2} \). The smoothness prior in [43] assumes small flow gradients over a 4-neighborhood of horizontal and vertical flow components. The non-local term encourages smoothness over a neighborhood of size \( 5 \times 5 \).

In both cases, the smoothness assumption is represented as

\[
f_{\text{S}} (x_{ij}, x_{i'j'}) = f_{\text{N}} (x_{ij}, x_{i'j'}) = x_{ij} - x_{i'j'} \tag{16}
\]

Akin to energy minimization approaches [43], we found the use of the non-local term crucial for obtaining accurate optical flow estimates, since mean-field inference using only the terms \( E_{\text{D}} \) and \( E_{\text{S}} \) is rather outlier-prone.

5.2. Probabilistic FlowFields

In a second setup, we aim for a probabilistic version of FlowFields [1]. We follow Bailer et al. and consider the energy used in the post-processing step of EpicFlow [36], which uses a data term based on gradient constancy. As proposed in [49], the image gradient terms are normalized w.r.t. the spatial derivatives, which is helpful to avoid outliers in the flow field. We obtain the linearized data term as

\[
f_{\text{D}}(y_{ij} ; I) = \sum_{r=1}^{3} \theta_{ij}^r \circ \left[ \nabla_{2} I_{2}^r (i + u_{ij}^{0}, j + v_{ij}^{0}) - \nabla_{1} I_{1}^r (i,j) + H (I_{2}^r (i + u_{ij}^{0}, j + v_{ij}^{0})) (y_{ij} - y_{ij}^{0}) \right]_{2}. \tag{17}
\]

Here, \( r = 1, \ldots, 3 \) indicates the RGB color channels, \( H (I_{2}^r) \) denotes the Hessian of \( I_{2}^r \), and \( \circ \) is the Hadamard product. The normalization coefficient \( \theta_{ij}^r \) is given as

\[
\theta_{ij}^r = \left( \frac{\theta_{ij,k}^{r_1}, \theta_{ij,k}^{r_2}}{2} \right), \quad \theta_{ij,k} = \frac{1}{\sqrt{\|H(I_{2})\cdot e_k\|_2^2 + \zeta^2}} \tag{18}
\]

with 2D unit vector \( e_k \) and a small constant \( \zeta > 0 \) [49].

The smoothness term of [36] is based on the flow gradient norm and uses additional filters of size \( 2 \times 3 \) and \( 3 \times 2 \) as described in [38] when calculating the flow derivatives. To keep the inference problem consistent, our approach differs slightly from [36] in that we only use the forward gradient filter described in Eq. (16) and obtain

\[
f_{\text{S}} (y_{ij}, y_{i'j'}) = \sqrt{(u_{ij} - u_{i'j'})^2 + (v_{ij} - v_{i'j'})^2} \tag{19}
\]
for \( y_{ijk} \) in a 4-neighborhood of \( y_{ij} \).

Following [36], a locally adaptive trade-off parameter \( \lambda_\Sigma(x) = \exp(-\kappa \| \nabla I_1(x) \|) \) is used. Similar to MAP estimation, the minimization of the KL divergence in Eq. (13) obtained from \( f_D, f_S \), and \( \lambda_\Sigma(x) \) performs well in practice. Therefore, an additional non-local term can be neglected.

6. Implementation

Optical flow estimation by energy minimization is known to be far from easy. Similarly, an application of mean-field inference is non-trivial in this context. Moreover, it is essential to consider several details commonly used with energy-based approaches to obtain satisfying results [15, 43]. A summary of basic design choices as well as an extensive analysis of the influence of all described specifics can be found in the supplemental material.

Common per-pixel updates, e.g. [48], do not work well for the mean-field inference of Eq. (13). Instead, we keep the corresponding optimization procedure as efficient as possible and use a block-coordinate descent scheme updating flow estimates \( \mu \), variances \( \Sigma \), and latent variables \( k \) in an alternating manner. We derive the gradient of the KL divergence in Eq. (13), set it to zero, and obtain an update equation for each set of variables (see supplemental). As with MAP, we found a joint update of the flow predictions at all pixels to be crucial to obtaining smooth flow fields.

Parameters. To determine suitable trade-off parameters \( \lambda \), it is not sufficient to follow the common approach and choose parameters that lead to the smallest AEPE on a training set [15, 43]. Instead, we evaluate the quality of both the obtained flow predictions and the uncertainty measure in order to ensure accurate flow estimates and meaningful entropies. To assess the quality of our uncertainty measure, we follow the approaches in [8, 24, 27, 30] and compute so-called sparsification plots. To that end, the pixels of a flow field are sorted according to the estimated uncertainties. Subsequently, an increasing percentage of the pixels is removed and the AEPE of the remaining pixels is calculated. In order to evaluate how well different uncertainty measures perform on an entire dataset, we propose to normalize the graphs, calculate the area under curve (AUC), and average over the sequences. To consider the trade-off between flow accuracy and quality of the uncertainty measure, we evaluate an \( F_1 \)-score

\[
F_1 = \frac{\text{AEPE} \cdot c \cdot \text{AUC}}{\text{AEPE} + c \cdot \text{AUC}}
\]

over the training data with constant \( c \) weighting the influence of the two metrics. We then determine parameters using Bayesian optimization [42] of Eq. (20).

Concerning the penalty functions, we follow the approach in [44] and learn appropriate GSM models for data likelihood, smoothness prior, as well as the non-local term from the respective training datasets or a randomly chosen subset thereof. Using manually determined variances \( \sigma_i \), a simple expectation maximization algorithm is used to obtain the corresponding weights \( \pi_i \). We observe that GSMS with \( L = 10 \) components perform well in practice. To save computational time, we resort to \( L = 5 \) components for our probabilistic implementation of Classic Flow.

Details. For Probabilistic Classic Flow, we follow the underlying energy approach (ClassicA, [43]) and use zero flow as an initialization; we denote the method as \( \text{ProbClassicA} \). Variances are initialized as \( \sigma_{\text{init}} = 1 - \frac{\epsilon}{7} \), latent variables as \( k_{\text{init}} = 1/L \). During the inference process, the parameter \( \lambda_C \) is annealed as in [43]. The parameters \( \lambda_D \), \( \lambda_S \), and \( \lambda_\Sigma \) are obtained with Bayesian optimization of the \( F_1 \)-score (Eq. (20)) using \( c = 5 \). As suggested by Sun et al. in the context of MAP, we use the variables \( \tilde{\mu} \) as the flow prediction. The corresponding uncertainties are then obtained from \( \Sigma \).

For Probabilistic FlowFields, we use the state-of-the-art FlowFields method [1] to generate sparse matches. Following Bailer et al., we apply the EpicFlow [36] post-processing step with Sintel parameters to interpolate the matches and use the results to initialize our algorithm. This method will be denoted as \( \text{ProbFlowFields} \). Variances and latent variables are initialized as for \( \text{ProbClassicA} \). Trade-off parameters \( \lambda_D \), \( \lambda_S \), and \( \kappa \) are obtained again with Bayesian optimization [42]. The parameter of the \( F_1 \)-score (Eq. (20)) is chosen as \( c = 100 \) since the AEPE is significantly higher on Sintel. We will make code available for \( \text{ProbClassicA} \) as well as for \( \text{ProbFlowFields} \).

7. Experiments & Results

In the following, we evaluate our probabilistic flow approach and assess the quality of our uncertainty measure by comparing it to different approaches from the literature.

7.1. Competing uncertainty measures

We apply existing uncertainty measures on top of the corresponding energy minimization approaches (i.e. ClassicA or FlowFields) in order to analyze the benefit of our combined flow prediction and uncertainty estimation. We limit ourselves to a few select methods here and give a more extensive comparison in the supplemental material.

Barron et al. [3] suggest a confidence measure based on the spatial gradient of the input image. The corresponding uncertainty is obtained as \( \Psi_{\text{Gradient}} = -\| \nabla_2 I_1 \| \) using central differences to approximate the gradient.

The learned confidence measure in [30] is based on a classifier that predicts with probability \( \tilde{p} \) whether the error of the flow estimate at a certain pixel is smaller than a specified threshold \( \epsilon_T \). We use the implementation of MacAodha et al. and train \( \Psi_{\text{Learned}} = -\tilde{p} \) as described in [30].
As proposed by Bruhn and Weickert [8], we use the local energy contribution of the MAP estimate $y_{\text{MAP}}$ as a model-inherent uncertainty $\Psi_{\text{Energy}} = E(y_{\text{MAP}}; I)$. We apply the non-linear version of the underlying energy function as it shows an improved performance [10]. A data penalty for out-of-boundary pixels is determined on the training set.

For an additional baseline, we perform a Laplace approximation of the posterior $p(y, \hat{y} \mid I)$ around $y_{\text{MAP}}$. With $H$ denoting the Hessian of the linearized energy $E(y_{\text{MAP}}; I)$, the covariance is given as $\Sigma_L = H^{-1}$. Similar to our approach, we obtain $\Psi_{\text{Laplace}} = -\log(\det(H)) + \text{const.}$

Finally, we consider an oracle uncertainty measure $\Psi_{\text{Oracle}}$, for which the uncertainty of a pixel is given by its end-point error. Thus, the estimate $\Psi_{\text{Oracle}}$ provides a bound for the best possible uncertainty estimation. Note that we consider oracle uncertainties from the predictions obtained by MAP estimation: the oracle uncertainties for our flow predictions show a very similar behavior.

### 7.2. ProbClassicA

We first evaluate the application of ProbFlow to the ClassicA model described in Sec. 5.1. We rely on the commonly used AEPE. To compare the performance of different flow estimation algorithms, we rely on the Middlebury dataset [2], which has frequently been used to evaluate the performance of optical flow confidence measures. In order to compare different uncertainty approaches, ground truth optical flow is needed. As the Middlebury training set includes only 8 image pairs, we need to resort to training and testing the uncertainty measures on the same data.

To compare the performance of different flow estimation algorithms, we rely on the commonly used AEPE. Table 1 gives results on Middlebury training and test. Our method (ProbClassicA) performs on par with ClassicA. On the one hand, this is to be expected as both approaches share the same energy formulation and energy-based methods use highly elaborate schemes. On the other hand, this is the first time that a fully probabilistic method achieves competitive results on a public benchmark. For completeness, we also include the related Classic++, which shows slightly better results as its median filtering step allows for a more effective outlier suppression than an additional nonlocal term [25]. The median filter cannot easily be applied to our case, but the results for ClassicA and ProbClassicA could be further improved by adding weights to the nonlocal term [43].

The evaluation of uncertainty measures is more challenging. Sparsification plots are commonly considered [8, 24, 27, 30]. To be able to compare the performance of different uncertainty measures over an entire dataset, we calculate the AUC as described in Sec. 6. However, as argued by Márquez-Valle et al. [31], sparsification plots do not allow to estimate how strongly an uncertainty estimate is related to the underlying pixel errors. To address this, we propose to compute the Spearman’s rank correlation coefficient (CC), which estimates how well the examined uncertainty values can be mapped onto the corresponding end-point errors using an arbitrary monotonic function.

Table 2 and the sparsification plots in Fig. 2 show that our uncertainty predictions are clearly superior in detecting the most reliable flow estimates. The gradient uncertainty has almost no ability to rank the pixels according to their accuracy. Hence, the simple consideration of input data appears insufficient. Similarly, the Laplace measure does not lead to satisfying results. $\Psi_{\text{Energy}}$ and $\Psi_{\text{Learned}}$ lead to very similar AUC results, whereby the learned uncertainty takes substantial time for training. Our method improves the AUC over previous ones by more than 6%. Moreover, the evaluation of the CC reveals that our uncertainty measure shows by far the highest correlation between the assigned uncertainty value and the per-pixel end-point error with a relative improvement of 13% over the learned uncertainty.

### 7.3. ProbFlowFields

Next, we apply our probabilistic approach to the competitive FlowFields method as described in Sec. 6. We use
the more recent Sintel benchmark [11], which in comparison to the Middlebury dataset, allows to partition its more than 1000 flow sequences into training and validation sets. Exemplary flow and uncertainty estimates as well as the corresponding ground truth can be seen in Fig. 3.

Table 3 summarizes the AEPEs on our validation set and the test set of the Sintel benchmark. For comparison, we also report results of FlowFields*, based on the same consistent EpicFlow energy variant underlying ProbFlowFields, c.f. Sec. 5.2. Again, our estimates from ProbFlowFields are competitive with its underlying energy method and we observe results on par with the original FlowFields.

Evaluated on the Sintel test set, ProbFlowFields currently ranks 6th in comparison to previously published methods. Please note that the FlowFields test results shown in Table 3 are superior to the publicly available AEPE of 5.810. To suppress the effect of the random component in the matches of FlowFields and thus have a fair comparison, we use FlowFields [1] to compute forward-backward (FB) flow for all consecutive image pairs. Subsequently, points are sampled on a regular grid in the first frame and tracked as long as (1) their FB flows are consistent, and (2) the gradient magnitude of the flow is below a threshold. If tracks are stopped, new points are inserted to preserve sampling regularity unless no points can be tracked in a region.

In a second setting, we apply ProbFlowFields using parameters trained on Sintel to compute FB flow along with normalized forward uncertainties \( \Psi^F \). Here, we keep the FB condition (1) and use a new condition (3) such that we end any track passing through location \((i, j)\) if the pixel uncertainty \( \Psi^F_{ij} \) falls below an empirically determined threshold.

Our results in terms of segmentation precision, recall, F-measure, as well as the achieved trajectory density on the FBMS-59 dataset [33] are given in Table 5. For both approaches, sparse trajectories are sampled at 8 pixel distance, and framewise dense segmentations are computed using the variational approach of [32]. Moreover, we compare to the current state-of-the-art [22] on FBMS-59, which is based on Large Displacement Optical Flow (LDOF) [7].

The evaluation of sparse segmentations shows that ProbFlow allows for a higher average point density of more

<table>
<thead>
<tr>
<th>Uncertainty measure</th>
<th>AUC</th>
<th>rel. chg.</th>
<th>CC</th>
<th>rel. chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gradient [3]</td>
<td>1.022</td>
<td>1.57</td>
<td>-0.009</td>
<td>1.02</td>
</tr>
<tr>
<td>Laplace</td>
<td>0.657</td>
<td>0.65</td>
<td>0.257</td>
<td>0.54</td>
</tr>
<tr>
<td>Energy [8]</td>
<td>0.470</td>
<td>0.18</td>
<td>0.434</td>
<td>0.23</td>
</tr>
<tr>
<td>Learned [30]</td>
<td>0.474</td>
<td>0.19</td>
<td>0.451</td>
<td>0.20</td>
</tr>
<tr>
<td>ProbFlowFields (ours)</td>
<td>0.398</td>
<td>0.00</td>
<td>0.563</td>
<td>0.00</td>
</tr>
<tr>
<td>Oracle</td>
<td>0.182</td>
<td>–</td>
<td>1.000</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 4. Area under curve (AUC), Spearman’s rank correlation coefficient (CC), and relative change (rel. chg.) in comparison to our uncertainty measure on a Sintel benchmark validation set.

7.4. Application to motion segmentation

We now show the benefit of our uncertainty estimates when optical flow is used as a cue for motion segmentation. Current state-of-the-art methods (e.g., [22, 33]) build upon precomputed point trajectories, i.e., spatio-temporal curves that describe the individual point motion over extended periods. Ideally, such point trajectories are sampled with a regular density, they are long and reliable [6].

We build on the minimum cost multicut approach of [22] to obtain sparse motion segmentations. We follow Keuper et al. and generate point trajectories as in [6]. In a first setup, we use FlowFields [1] to compute forward-backward (FB) flow for all consecutive image pairs. Subsequently, points are sampled on a regular grid in the first frame and tracked as long as (1) their FB flows are consistent, and (2) the gradient magnitude of the flow is below a threshold. If tracks are stopped, new points are inserted to preserve sampling regularity unless no points can be tracked in a region.

In a second setting, we apply ProbFlow using parameters trained on Sintel to compute FB flow along with normalized forward uncertainties \( \Psi^F \). Here, we keep the FB condition (1) and use a new condition (3) such that we end any track passing through location \((i, j)\) if the pixel uncertainty \( \Psi^F_{ij} \) falls below an empirically determined threshold.

Our results in terms of segmentation precision, recall, F-measure, as well as the achieved trajectory density on the FBMS-59 dataset [33] are given in Table 5. For both approaches, sparse trajectories are sampled at 8 pixel distance, and framewise dense segmentations are computed using the variational approach of [32]. Moreover, we compare to the current state-of-the-art [22] on FBMS-59, which is based on Large Displacement Optical Flow (LDOF) [7].

The evaluation of sparse segmentations shows that ProbFlow allows for a higher average point density of more

<table>
<thead>
<tr>
<th>Uncertainty measure</th>
<th>AUC</th>
<th>rel. chg.</th>
<th>CC</th>
<th>rel. chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProbFlowFields</td>
<td>3.127</td>
<td>0.00</td>
<td>5.696</td>
<td>0.00</td>
</tr>
<tr>
<td>ProbFlowFields + BS</td>
<td>3.052</td>
<td>-0.02</td>
<td>5.628</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

Table 3. Average end-point error (AEPE) and relative change (rel. chg.) w.r.t. to ProbFlowFields on Sintel. *FlowFields shows better results than the ones published on the website. See text for details.
than 1.18% compared to a maximal density of 0.89% with FlowFields [1] or LDOF [7]. Low point density usually indicates flow inconsistencies in homogeneous regions and therefore a more uneven sampling. As can be seen by the increased F-measure on sparse and especially on densified segmentations, the improvement in the trajectory computation directly translates to an improved segmentation quality.

To illustrate the complementarity of our uncertainties to the FB condition, we perform an experiment using ProbFlowFields estimates and the FB check (1), but not the corresponding uncertainties, i.e. omitting (3). In this case, the F-measure of the sparse matches drops significantly. Using FlowFields estimates and the post-hoc measure $\Psi_{\text{Energy}}$, the uncertainties again complement (1). However, the usage of ProbFlowFields yields a higher F-measure, thus showing a clear benefit also in this application.

8. Conclusion

To address the issue that optical flow estimates are not equally reliable throughout an image, we introduced ProbFlow – a probabilistic framework for the joint estimation of optical flow and its underlying uncertainty. Starting from conventional energy minimization methods, we derived the posterior distribution and used variational inference to estimate the flow and a model-inherent uncertainty. This is unlike existing uncertainty measures that detect regions of unreliable flow in a post-hoc step. We applied our approach to two different energy formulations on the Middlebury and Sintel benchmarks, where we obtain competitive flow estimates and significantly improved uncertainties. Applying our uncertainty estimates in the context of motion segmentation, we were able to discard erroneous flow estimates and generate highly reliable point trajectories.

Acknowledgments. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007–2013)/ERC Grant agreement No. 307942. We would like to thank Tobias Plötz and Jochen Gast for helpful discussions.
References


[22] M. Keuper, B. Andres, and T. Brox. Motion trajectory segmentation via minimum cost multicut. ICCV, 2015. 2, 7, 8


[38] S. Roth and M. J. Black. Steerable random fields. ICCV, 2007. 4


