2017 IEEE International Conference on Robotics and Automation (ICRA)
Singapore, May 29 - June 3, 2017

Accurate Stereo Visual Odometry with Gamma Distributions

Ruben Gomez-Ojeda, Francisco-Angel Moreno, Javier Gonzalez-Jimenez

***Abstract*— Point-based stereo visual odometry systems typically estimate the camera motion by minimizing a cost function of the projection residuals between consecutive frames. Under some mild assumptions, such minimization is equivalent to maximizing the probability of the measured residuals given a certain pose change, for which a suitable model of the error distribution (sensor model) becomes of capital importance in order to obtain accurate results. This paper proposes a robust probabilistic model for projection errors, based on real world data. For that, we argue that projection distances follow Gamma distributions, and hence, the introduction of these models in a probabilistic formulation of the motion estimation process increases both precision and accuracy. Our approach has been validated through a series of experiments with both synthetic and real data, revealing an improvement in accuracy while not increasing the computational burden.**

I. Introduction

Most stereo visual odometry systems estimate camera motion through the least-squares minimization [1], [2] of a certain cost function $C(\boldsymbol{\xi})$ of the residuals $\Delta\mathbf{p}_i(\boldsymbol{\xi})$, defined as difference between the *observations* $\mathbf{p}'_i$ of a set of keypoints and their *predictions* $\hat{\mathbf{p}}_i(\boldsymbol{\xi})$ [3]:

$$\Delta\mathbf{p}_i(\boldsymbol{\xi}) = \hat{\mathbf{p}}_i(\boldsymbol{\xi}) - \mathbf{p}'_i, \tag{1}$$

where $\hat{\mathbf{p}}_i(\boldsymbol{\xi})$ is computed by back-projecting to 3D the observed i-th image keypoint detected in the previous frame, and then re-projecting it to the current one, according to an estimation of the pose change $\boldsymbol{\xi} \in \mathfrak{se}(3)$ between them. Typically, the cost function is derived from the maximization of the probability of the pose change given the residuals, so that minimizing $C(\boldsymbol{\xi})$ (i.e. estimating the optimal pose change $\boldsymbol{\xi}^*$) is equivalent to maximizing $p(\boldsymbol{\xi}\,|\,\Delta\mathbf{p})$. This is also equivalent to maximizing their likelihood given a certain pose change, under the assumptions of independent and equally distributed noise, and a uniform *prior* distribution over the poses:

$$\boldsymbol{\xi}^* = \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\, p(\boldsymbol{\xi} \mid \Delta\mathbf{p}) = \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\, p(\Delta\mathbf{p} \mid \boldsymbol{\xi}). \tag{2}$$

In this context, finding a proper model of the residual distribution becomes of capital importance as the results will be directly affected by its goodness of fit. Furthermore, such model should consider the presence of not only noise but also outliers.

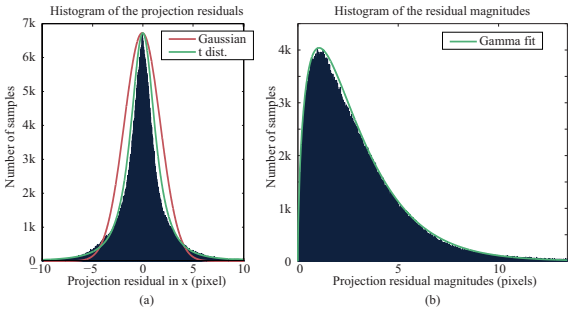

Fig. 1. Histogram of (left) the projection residuals in the x-coordinate for the keypoints extracted from the sequence KITTI00, and the corresponding fitted Student's t-distribution (in green) and Gaussian distribution (in red), and (right) the residual magnitudes of the keypoints extracted from all the training sequences in the KITTI dataset. The real distribution can be described accurately with a Gamma distribution.

Commonly, keypoint predictions (and consequently the projection residuals) are considered to be Gaussian-distributed, since the observed keypoints are assumed to be corrupted by Gaussian noise that is propagated through linear approximations of the above-mentioned back-projection and re-projection functions [4]. In practice, though, such approximations still present inaccuracies, as can be seen in Figure 1(a), which depicts the distribution of the residuals in the image x-coordinate computed from real data (which is similar to that of the y-coordinate). As a consequence, assuming a Gaussian distribution for $p(\Delta\mathbf{p}|\boldsymbol{\xi})$ leads to an unsuitable cost function whose minimization will yield inaccurate results. In fact, according to the real distribution, a better approach would be to model the residual in both the x and y image coordinates by a Student's t-distribution, whose shape is similar to the Gaussian one but with heavier tails (refer again to Figure 1(a)). This approach has been explored in [5] and applied to RGB-D cameras.

Nevertheless, we propose to employ the *magnitude* of the residual between observations and predictions $\mathbf{r} = \{r_i(\boldsymbol{\xi}) = \|\Delta\mathbf{p}_i(\boldsymbol{\xi})\|_2\}$, instead of the projection residual $\Delta\mathbf{p}$. Thus, we claim that modeling $\mathbf{r}$ as a Gamma distribution (i.e. $\mathbf{r} \sim \Gamma(\theta, \alpha)$) is a better option than modeling residuals as either a Gaussian or a t-distribution, since the fitted model deviate less from the actual distribution it approximates (see Figure 1(b)). Following this, and the above-mentioned assumptions, the optimization problem

MAPIR Group. Universidad de Málaga. E.T.S. Ingeniería de Informática-Telecomunicación. Campus de Teatinos, 29071, Málaga, Spain. E-mail: rubengooj@gmail.es

This work has been supported by the project "PROMOVE: Advances in mobile robotics for promoting independent life of elders", funded by the Spanish Government and the "European Regional Development Fund ERDF" under contract DPI2014-55826-R.

in (2) becomes:

$$\boldsymbol{\xi}^* = \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\, p(\boldsymbol{\xi} \mid \mathbf{r}) = \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\, p(\mathbf{r} \mid \boldsymbol{\xi}). \tag{3}$$

The introduction of the Gamma distribution in the optimization process allows us to derive a more suitable cost function that leads to better results than assuming that residuals follow either a Gaussian or a t-distribution, as will be proved with a series of experiments.

The two parameters of such Gamma distribution (namely shape and scale) are estimated at each time-step from the actual histogram of all the involved residual magnitudes, being necessary a minimum number of samples for the fit to be representative. The on-line fitting procedure introduces little additional cost to the optimization process while the benefits are two-fold: a more precise camera pose estimation and more robustness against outliers and noisy measurements than the standard Gaussian-based approach. Even so, a very large ratio of outliers may eventually degrade its performance, so the usage of *robustification* methods is still advisable.

Our claim is supported by an extensive experimental validation with both synthetic and real data, revealing its suitability for performing visual odometry, specially for stereo vision systems where observations in both images can be employed. For that, our proposal has been integrated into our previous stereo visual odometry (SVO) system presented in [6]. The results show significant improvements in accuracy whilst incurring in a reduced computational footprint. An illustrative video of our system and the SVO library source code can be accessed in `http://mapir.uma.es/rgomez`.

II. Related Work

Visual-based motion estimation algorithms are strongly affected by the presence of noisy data and, specially, outliers, which do not follow the commonly assumed Gaussian distribution for the residuals, hence eventually leading the system to erroneous results. Traditional approaches to this problem, as the one in [7], often rely on variants of RANSAC to deal with wrong measurements by generating a solution which is in consensus with the majority of the dataset. However, this technique has high computational requirements. In [8], Person et al. presented a stereo visual odometry system which takes advantage of monocular techniques, as they argue that those techniques are more refined and robust than those of stereo systems. For that, they implement a delayed outlier identification procedure based on an essential matrix RANSAC approach and robust iterative triangulation. Other approaches, as the one in [9], integrate robust probabilistic filters to explicitly deal with outliers by estimating, for instance, the depth at feature locations over multiple frames. Then, these depth filters are updated at each frame labeling as inliers those points with low uncertainty in depth, hence being introduced into the map and subsequently employed to estimate the camera motion.

Finally, other approaches introduce some robust cost functions in the camera motion estimation, hence obtaining appropriate weights that reduces the impact of wrong measurements. A first group proposes several modifications of the well-known extended Kalman filter (EKF) in order to increase the robustness of their systems against outliers and noisy measurements. In [10] authors propose a robust EKF filter to deal with outliers in real-time, by down-weighting the samples with more probability of being outliers, for which they learn the system dynamics thus avoiding manual parameter tuning. In [11] the previous approach was generalized and extended by introducing efficient smoothing and filtering modifications for dealing with data corrupted with non-Gaussian and heavy-tailed noise. The previous work was also extended in [12], where authors proposed to introduce a structured variational approximation with a more robust and flexible behavior, and yet introducing only a little increment in the computational complexity. Another group of techniques model directly the error distribution, and then perform a robust non-linear least-squares minimization of these errors. In [5], Kerl et al. perform robust odometry estimation for RGB-D cameras by minimizing the photometric error between two consecutive frames. They argue that their dense RGB-D residuals can be better explained with Student's t-distributions, for which they derive a probabilistic formulation including a robust sensor model based on real world data. Recently, the work in [13] proposes a generic self-tuning M-estimator which iteratively estimates the parameters of the residual distribution, thus removing the necessity of manually set such parameters. However, this method needs to compute the importance weights for each iteration of the least-squares problem, hence being computationally expensive for small problems as the one we address here.

III. Distribution of the Projection Errors and Residuals

In this section, we empirically analyze the actual distributions of both the projection residual $p(\Delta\mathbf{p} \mid \boldsymbol{\xi})$ and the residual magnitude $p(\mathbf{r} \mid \boldsymbol{\xi})$ for the case of image keypoints. For this purpose, we detect and match ORB keypoints along a sequence of stereo images provided by the KITTI collections of public datasets [14]. Then, the observed keypoints are projected to the next frame by applying the ground truth pose increment (also included in the dataset), and both the residuals and their magnitude are computed. Finally, we adjust different distributions to the data and evaluate their goodness of fit.

Regarding the projection residuals, we refer again to Figure 1(a), which has been built from the sequence "00" of the KITTI dataset. As stated before, it can be seen that the residual in the x image coordinate (and similarly for the y coordinate) does not follow a Gaussian distribution. In fact, these data can be more properly fitted by a t-distribution, as pointed out in [5].

TABLE I

AVERAGE GOODNESS OF FIT WITH THE K-S TEST FOR EACH DISTRIBUTION, WITH A CRITICAL VALUE OF 0.0608 FOR $\alpha = 0.05$

			Proj. Residual		Magnitude
Seq.	Frames	Feats	Gaussian	t-dist.	Gamma
00	4540	777k	0.1455	0.0827	**0.0474**
01	1100	100k	0.2327	0.1447	**0.0591**
02	4660	1059k	0.1035	0.0638	**0.0474**
03	800	200k	0.1222	0.1518	**0.0475**
04	270	55k	**0.0533**	0.1199	**0.0532**
05	2760	478k	0.1023	0.0725	**0.0492**
06	1100	133k	0.1673	0.1256	**0.0499**
07	1100	199k	0.1790	0.1313	**0.0464**
08	4070	731k	0.1412	0.0916	**0.0456**
09	1590	273k	0.1429	0.0745	**0.0481**
10	1200	191k	0.1298	**0.0534**	**0.0478**

So, we may consider to use this distribution to derive a suitable cost function that takes into account a better approximation of the residual true distribution. However, we claim that modeling the residual magnitude as a Gamma distribution instead of the residual as either a Gaussian or a t-student represents a more accurate fit of the modeled variable.

To prove this, we also analyze the distribution of the residual magnitude, shown in Figure 1(b), where all the training sequences in the KITTI dataset have been employed to build the histogram. It can be observed that a Gamma distribution accurately describes the behavior of the magnitude, as it presents a certain bias and also a heavy tail. The goodness of the three fits (i.e. Gaussian and t-distribution for the projection residual and Gamma for the residual magnitude) are evaluated through the Kolmogorov-Smirnov (K-S) test [15], which measures the maximal difference between an empirical and a real distribution function. Thus, for each sequence, a subset of 10^3 keypoints has been randomly selected from all the found features so that half of them are employed to derive the distribution model, while the rest is used to perform the test. Note that using separate datasets is mandatory in order to obtain valid and distribution-free K-S test results [16], hence allowing the comparison of different distributions. This experiment has been repeated 10^3 times for each sequence, obtaining the average values shown in Table I. In all sequences, the values below the test's critical value have been highlighted (which is 0.0608 for a significance value of $\alpha = 0.05$).

As expected, the t-distribution approach approximates better the real distribution of the residual than the Gaussian model. Nonetheless, the results also reveal that the Gamma distribution represents a more accurate model for the residual magnitude than the t-distribution for the projection residual in most datasets. Then, modeling their magnitude as a Gamma distribution (and consequently deriving a cost function of the residual magnitude according to that) will lead to more accurate results than employing a cost function of the projection residuals based on the t-distribution.

Finally, it is important to remark that the number of samples (i.e. observed keypoints) employed to fit the distributions influences the quality of the approximation, as will be further discussed in Section V.

IV. MOTION ESTIMATION WITH THE GAMMA DISTRIBUTION

In this section, we derive the equations to robustly recover the 6D pose change $\boldsymbol{\xi}$ of a stereo camera using the Gamma-based approach to model the behavior of the residual magnitude. For that, let us formally define the vector of residual magnitudes $\mathbf{r}(\boldsymbol{\xi}) = \{r_i(\boldsymbol{\xi})\}$ that contains the projection distances of all the individual observations, as defined in Section I. Then, we aim to find the camera motion $\boldsymbol{\xi}^* \in \mathfrak{se}(3)$ that maximizes the posterior probability $p(\boldsymbol{\xi} \mid \mathbf{r})$ as stated in equation (3), which we reproduce here for clarity:

$$\boldsymbol{\xi}^* = \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\, p(\boldsymbol{\xi} \mid \mathbf{r}) = \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\, p(\mathbf{r} \mid \boldsymbol{\xi}). \tag{4}$$

Under the mild assumptions of $r_i(\boldsymbol{\xi})$ being independent, estimating (4) is equivalent to minimizing the negative log-likelihood of the residual magnitude (refer to [17], [18] for further details):

$$\boldsymbol{\xi}^* = \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\, p(\mathbf{r} \mid \boldsymbol{\xi}) = \underset{\boldsymbol{\xi}}{\operatorname{argmin}} \Big\{ -\sum_i log\, p(r_i \mid \boldsymbol{\xi}) \Big\} \tag{5}$$

Now, we model the magnitude $\mathbf{r}(\boldsymbol{\xi})$ with a Gamma distribution, i.e. $\mathbf{r} \sim \Gamma(\alpha, \theta)$, whose probability density function (pdf) is given by:

$$f\big(x; \alpha, \theta\big) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-x/\theta} \;\text{ for } x > 0 \text{ and } \alpha, \theta > 0 \tag{6}$$

where α and θ are the so-called shape and scale parameters, respectively. Then, the individual likelihood of the residual magnitude is proportional to:

$$p(r_i \mid \boldsymbol{\xi}) \propto r_i^{\alpha-1} e^{-r_i/\theta} \tag{7}$$

where we have dropped the constant terms that do not depend on $\boldsymbol{\xi}$. Finally, by introducing this model into (5), the estimator becomes:

$$\boldsymbol{\xi}^* = \underset{\boldsymbol{\xi}}{\operatorname{argmin}} \sum_i \Big\{ r_i/\theta - (\alpha - 1) log\, r_i \Big\} \tag{8}$$

which is equivalent to minimizing this cost function (following an Iteratively Re-Weighted Least Squares (IRLS) approach):

$$\boldsymbol{\xi}^* = \underset{\boldsymbol{\xi}}{\operatorname{argmin}} \sum_i w\big(r_i(\boldsymbol{\xi})\big) r_i^2(\boldsymbol{\xi}) \tag{9}$$

with $w\big(r_i(\boldsymbol{\xi})\big)$ being a weighting function defined by:

$$w\big(r_i(\boldsymbol{\xi})\big) = \frac{r_i/\theta - (\alpha - 1) log\, r_i}{r_i^2}. \tag{10}$$

For the sake of computational complexity, we fit the Gamma distribution at each time-step with the Method of Momentums, which employs the closed form solutions

for both the mean $\mu = \alpha\theta$, and the variance of the distribution $\sigma^2 = \alpha\theta^2$ (a complete comparative of several methods for fitting Gamma distributions can be found in [19]). However, using these parameters entails that outliers also have influence in the estimation of both the mean and variance. Therefore, it is desirable to employ robust methods to estimate the distribution parameters [20]. Thus, we employ the Median Absolute Deviation for the standard deviation $\hat{\sigma}$ and, subsequently, we estimate a robust mean $\hat{\mu}$ by only considering the samples lying less than three times $\hat{\sigma}$.

Once the Gamma distribution has been fitted, we derive the weighted cost function in equation (9), which is optimized on the $\mathfrak{se}(3)$ manifold through the well-known Gauss-Newton equations (refer to [21] for a thorough analysis on on-manifold optimizations). Finally, although our proposed cost function presents some robustness against outliers due to the sublinear nature in the residual magnitude of the weighting equation (10), it is important to remark that a big amount of them still may degrade the resulting ego-motion estimation. Therefore, we have implemented a variant of the ERODE outlier detector [22] with the main difference of using a Cauchy distribution instead of a Huber loss function and employed as an outlier-removal strategy for all the tested approaches in the experiments.

V. Experimental Evaluation

This section presents two sets of experiments that analyze the effect in the localization accuracy of introducing the proposed Gamma-based model in our robust SVO system [6] in comparison to other approaches, namely: i) non-weighted, ii) Gaussian-weighted, and iii) Student's t-distributed weighted. For the first approach, we perform a standard least-squares minimization of the residuals without defining any weight for them. For the last two, we fit a Gaussian or a t-distribution to the computed projection residuals, respectively, and derive a cost function from equation (5), which will define the weights for the individual residuals.

A. Experiments with synthetic data

In this first set of experiments, we have generated random stereo observations (keypoints) in two consecutive frames, related by a random camera motion. Thus, image keypoints are randomly spread all over the first stereo pair, by simulating the point locations in the left image as well as their corresponding disparities. Then, we project them to the current stereo frame according to a random camera motion and, subsequently, Gaussian distributed noise is added to each keypoint in both stereo frames. Finally, we compute the motion estimation error in different scenarios.

In these experiments, we have simulated camera motions that follow a uniform distribution between ± 1 m and ± 3 deg, which emulates a camera moving at similar speeds to those presented in [14]. The disparity of stereo points has been set to follow a uniform distribution between $[10, 30]$ pixels, while the camera intrinsic parameters are those specified for the KITTI dataset.

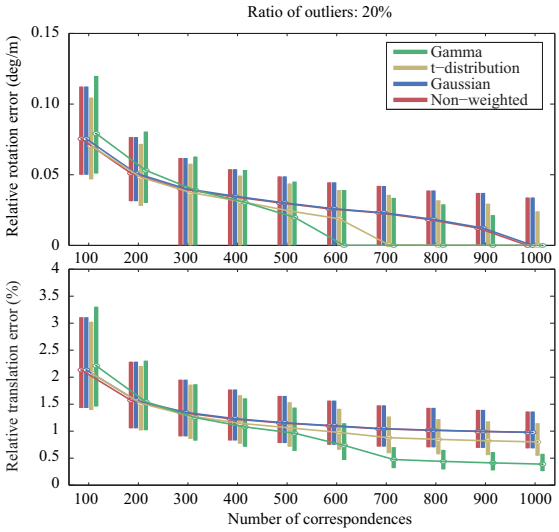

Fig. 2. Rotation (top) and translation (bottom) errors over a variable number of observations, employing different cost functions: non-weighted (in red), Gaussian weighted (in blue), Student's t-distribution (in yellow), and Gamma distribution (in green).

1) Impact of the Number of Observations: As discussed in Section I, the number of keypoint correspondences has a strong influence in the quality of the fitted Gamma distribution. To assess this, we have evaluated our SVO approach for a variable number of observations through the following Monte-Carlo simulation: for each weighting method and number of observations, we estimate the camera pose change for 1000 different configurations of both observations and camera motions, resulting in 1 million simulations. The outliers ratio has been set to 20 % in this series of experiments.

Figure 2 plots the results for the evaluated methods, where we have measured both rotation and translation average errors (along with 95% confidence intervals, plotted as solid bars), specified in deg/m and % of the total length, respectively, with respect to the true camera motion. As expected, both rotation and translation errors show a slightly superior performance of the other three methods in comparison to our approach for the lowest number of observations, since there is not enough information to fit a Gamma distribution properly. In contrast, this tendency is inverted as the number of observations increases, revealing our method to clearly outperform the other three approaches in both precision and accuracy, specially over 600 landmarks.

2) Impact of the Ratio of Outliers: Now, we study the impact of the number of outliers in the accuracy of the camera motion estimation, keeping a fixed number of 200 keypoints. Again, we have performed 1 million simulations for 1000 different configurations of observations and camera motions, respectively.

The results are depicted in Figure 3, where it is

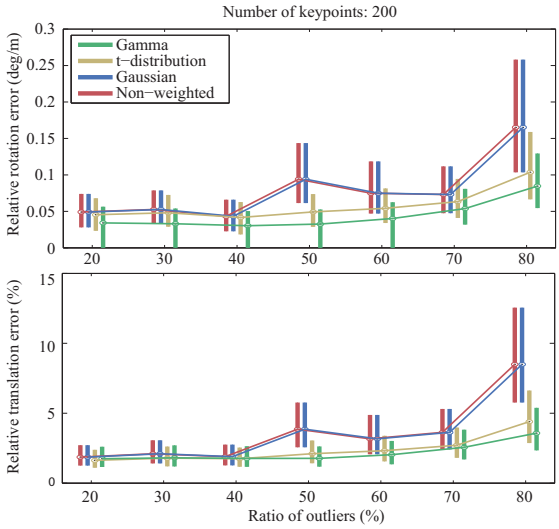

Fig. 3. Rotation (top) and translation (bottom) errors over all sub-sequences of a given length in the KITTI dataset, employing different cost functions: non-weighted (in red), Gaussian weighted (in blue), Student's t-distribution (in yellow), and Gamma distribution (in green).

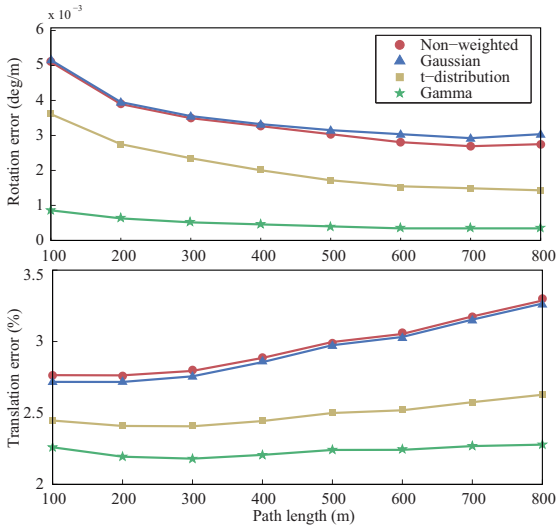

Fig. 4. Average rotation (top) and translation (bottom) errors over all sub-sequences of a given length in the KITTI dataset, employing different cost functions: non-weighted (in red circles), Gaussian weighted (in blue triangles), Student's t-distribution (in yellow squares), and Gamma distribution (in green stars).

shown a clearly better performance of both Gamma and t-distribution against the non-weighted and Gaussian-weighted approaches, since the first ones present a robust behavior (as discussed in previous section). Gamma and t-distribution approaches performs similarly in both translation and rotation errors for lower ratios of outliers, although the Gamma-based sensor model provide more accurate results when increasing the number of outliers, since they can be easily detected and removed from the residual magnitude distribution.

B. SVO Evaluation in the KITTI Benchmark

In this section, we assess the impact of the different approaches when performing robust camera ego-motion estimation for the training sequences ("00" to "10") from the KITTI dataset [14]. For that purpose, we have evaluated the results by using again the same metrics employed in the KITTI Benchmark, which computes errors in both rotation and translation for different sub-sequences lengths and speeds.

1) Performance at Different Sequence Lengths: First, we compute both rotation and translation errors relative to the distance traversed, for all the different subsequence lengths considered in the dataset (100m, 200m,..., 800m). The results show a significant improvement in both errors for all the different subsequences with our method, which performs clearly better than the rest of the approaches (refer to Figure 4).

Although the t-distributed weighting scheme also improves the ego-motion estimation in comparison to the Gaussian-weighted approach, our proposal yields better results since it describes more accurately the actual nature of the residual magnitude distribution. Moreover, it can be seen that the relative improvement of our approach, in comparison with the rest, grows as the path length increases. This is caused, in part, by the good performance obtained regarding to rotations, since high errors in rotation deviate the absolute trajectory from the ground truth, hence increasing absolute translational errors.

2) Performance at Different Speeds: In this experiment, we analyze the impact of the different weighting functions when performing visual odometry for all the speeds considered in the KITTI Benchmark (4m/s,6m/s,...,24m/s), by computing again the average rotation and translation errors (refer to Figure 5). It can be seen that our proposal clearly presents a superior performance for all the considered speeds, specially over 60km/h. This is caused by an increasing number of wrong measurements and outliers introduced to the system when the camera is traveling at high speeds, due to difficulties in feature tracking (those sequences usually correspond to low-textured highway scenes). In these situations, our proposed Gamma-based model performs better than the rest of the methods since it describes the actual nature of the residual magnitude distribution, so that outliers and wrong measurements are down-weighted properly, as claimed in this paper.

3) Computational Time: Finally, we analyze the computation time employed by each algorithm in the optimization process, for all the frames in the training set of the KITTI dataset. Experiments have been conducted on a single core of an Intel(R) Core(TM) i7-3770 CPU @ 3.40GHz processor with 4GB RAM. Table II contains the average time per frame employed for each algorithm, for a given number of observations. As expected, the non-weighted approach has the lowest computational footprint, as it

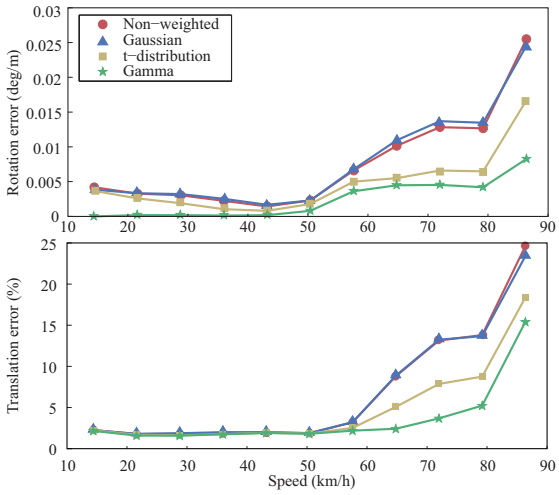

Fig. 5. Average rotation (top) and translation (bottom) errors over all sub-sequences of a given speed in the KITTI dataset, for the different cost functions considered.

TABLE II

AVERAGE OPTIMIZATION TIME PER FRAME FOR A GIVEN NUMBER OF OBSERVATIONS.

#Observations	nWeight.	Gauss.	t-dist.	Gamma
$N \leq 200$	0.911 ms	1.034 ms	1.134 ms	1.113 ms
$200 < N \leq 300$	1.399 ms	1.625 ms	1.783 ms	1.748 ms
$300 < N \leq 400$	1.872 ms	2.212 ms	2.420 ms	2.378 ms
$400 < N \leq 500$	2.329 ms	2.787 ms	3.038 ms	2.991 ms
$N > 500$	2.962 ms	3.590 ms	3.899 ms	3.854 ms

does not involve any weight estimation. On the other hand, although the Gaussian weighted approach requires less computational cost than the Gamma-weighted and t-distributed weighted approaches, it performs similar to the non-weighted approach (as demonstrated in the previous experiments), thus not justifying its application. Finally, our proposal slightly outperforms the t-distributed weighting scheme, with a smaller computational burden, while increasing the accuracy in visual odometry estimation. Hence, it is interesting to consider the inclusion of the proposed Gamma-based model in robust systems with higher requirements in accuracy than in computational time.

VI. CONCLUSIONS

In this paper we have proposed a Gamma-based model for the distribution of the projection residual magnitude in keypoint-based stereo-visual odometry. This approach is employed to derive a proper cost function of the residual magnitude which accurately weights each individual observation according to their true distribution. Its minimization leads to a robust ego-motion estimation that outperforms other weighting approaches that model projection errors as Gaussian or Student's t-distributions. Moreover, our proposal also presents robustness against outliers, since the model reproduces the tail behavior of the residual magnitude real distribution so that outliers are properly down-weighted in the optimization process. The claimed features have been proved with extensive visual odometry experiments with both synthetic and real data, where we compare our approach with the non-weighted, Gaussian-distributed, and t-distributed approaches.

REFERENCES

[1] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping," *IEEE Robotics & Automation Magazine*, vol. 13, no. 2, pp. 99–116, 2006.

[2] F. Fraundorfer and D. Scaramuzza, "Visual odometry: Part ii: Matching, robustness, optimization, and applications," *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, pp. 78–90, 2012.

[3] A. Gelb, *Applied optimal estimation*. MIT press, 1974.

[4] J. C. Clarke, "Modelling uncertainty: A primer," *Tutorial of Department of Eng. Science*, pp. 1–21, 1998.

[5] C. Kerl, J. Sturm, and D. Cremers, "Robust odometry estimation for rgb-d cameras," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pp. 3748–3754, IEEE, 2013.

[6] R. Gomez-Ojeda and J. Gonzalez-Jimenez, "Robust stereo visual odometry through a probabilistic combination of points and line segments," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*, IEEE, 2016.

[7] B. Kitt, A. Geiger, and H. Lategahn, "Visual odometry based on stereo image sequences with ransac-based outlier rejection scheme," in *Intelligent Vehicles Symposium (IV), 2010 IEEE*, pp. 486–492, IEEE, 2010.

[8] M. Persson, T. Piccini, M. Felsberg, and R. Mester, "Robust Stereo Visual Odometry from Monocular Techniques," *IEEE Intelligent Vehicles Symposium*, no. Iv, pp. 686–691, 2015.

[9] C. Forster, M. Pizzoli, and D. Scaramuzza, "SVO : Fast Semi-Direct Monocular Visual Odometry," *IEEE International Conference on Robotics and Automation*, 2014.

[10] J.-A. Ting, E. Theodorou, and S. Schaal, "Learning an outlier-robust kalman filter," in *European Conference on Machine Learning*, pp. 748–756, Springer, 2007.

[11] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "An outlier-robust kalman filter," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pp. 1551–1558, IEEE, 2011.

[12] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "Approximate inference in state-space models with heavy-tailed noise," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5024–5037, 2012.

[13] G. Agamennoni, P. Furgale, and R. Siegwart, "Self-tuning M-estimators," pp. 4628–4635, 2015.

[14] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 3354–3361, IEEE, 2012.

[15] M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *Journal of the American statistical Association*, vol. 69, no. 347, pp. 730–737, 1974.

[16] G. J. Babu and C. Rao, "Goodness-of-fit tests when parameters are estimated," *Sankhyā: The Indian Journal of Statistics*, pp. 63–74, 2004.

[17] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[18] C. Kerl, *Odometry from rgb-d cameras for autonomous quadrocopters*. PhD thesis, Citeseer, 2012.

[19] B. R. Clarke, P. L. McKinnon, and G. Riley, "A fast robust method for fitting gamma distributions," *Statistical Papers*, vol. 53, no. 4, pp. 1001–1014, 2011.

[20] P. J. Huber, *Robust statistics*. Springer, 2011.

[21] J.-l. Blanco, "A tutorial on SE (3) transformation parameterizations and on-manifold optimization," no. 3, 2013.

[22] F.-A. Moreno, J.-L. Blanco, and J. González-Jiménez, "ERODE: An efficient and robust outlier detector and its application to stereovisual odometry," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pp. 4691–4697, IEEE, 2013.