Noise Estimation and Adaptive Filtering During Visual Tracking

Ibrahima J. Ndiour and Patricio A. Vela
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

***Abstract*— This paper proposes a procedure to characterize segmentation-based visual tracking performance with respect to imaging noise. It identifies how imaging noise affects the target segmentation as measured through local shape metrics (Sobolev and Laplace metrics). Such a procedure would be an important calibration step prior to implementing a visual tracking filter for a given need. We utilize the Bhattacharyya coefficient between the target and background intensity distributions to estimate the segmentation error. An empirical study is conducted to establish a correspondence between the Bhattacharyya coefficient and the segmentation error. The correspondence is used to adaptively filter temporally correlated segmentations. Preliminary results show improved performance when compared to fixed gains.**

Index Terms – Contour tracking, observers, contrast parameter, shape metrics.

I. Introduction

This paper considers the problem of achieving accurate segmentation-based tracking in the face of imaging noise. Obtaining noise models is an essential step in understanding how to deal with such issues. For image segmentation needs, noise models lead to analytical methods for optimizing segmentation strategies [9], [10].

Alternatively, algorithmic improvements may be proposed. There are a few main classes of such modifications. One involves improving the mathematical model and the discrimination energies associated with the segmentation process. The second involves imposing shape constraints on the target model [8]. These classes of improvements have been shown to also improve associated segmentation-based tracking algorithms [2]. A third class, uniquely suited to video, is to introduce filtering schemes [3].

The spatio-temporal correlation between video frames should provide sufficient information to remedy poor segmentations arising from imaging noise that cannot be handled by optimizing the individual segmentations. This paper provides an analysis of noise on the segmentation procedure and determines its effects utilizing curve comparison metrics [11], [14]. With the expected error rates, it is possible to demand a corrective gain for handling the expected segmentation error arising from imaging noise. In order to properly incorporate the expected error into the filtering procedure, the functional relationship between the image data and the segmentation quality must be ascertained.

The principal contributions of the work include: a methodology for utilizing a proven contrast parameter to derive expected segmentation errors that are geometrically relevant, an empirical procedure for identifying the optimal filter gain given the measured contrast, and the use of the optimal gain for probabilistic shape filtering.

Organization: Section 2 covers the contrast parameter, the shape metrics, and the characterization procedure. The filtering strategy and its relation to the error estimates are discussed in Section 3. A proposed adaptive shape filter is tested and compared against ground truth. Section 4 concludes the paper.

II. Quantification of segmentation error through a contrast parameter

Suppose that the target is a single connected object in the image to process. Let P_{in} and P_{out} be the intensity probability distribution functions (pdfs) of pixels inside and outside the object, respectively. Local to the object, an algorithm's ability to segment is directly related to the interior and exterior pdfs. Segmentation ability is related to how distinct the distributions P_{in} and P_{out} are (see Figure 1). When there is significant overlap between the target and background distributions, the segmentation is prone to errors. Conversely, when the distributions are distinct, the segmentation is reliable.

1) Distance between pdf's: The Bhattacharyya coefficient between two distributions p and q is defined as

$$\beta(p,q) = \int \sqrt{P_{in}(x) \cdot P_{out}(x)}\, dx.$$

It is a similarity measure between pdf's that varies in the range $[0\,,\,1]$. High values of β indicate overlapping pdf's (and suspect segmentations), while low values indicate distinct pdf's (and reliable segmentations).

2) Distances between curves: Several metrics [1] exist to quantify the result of a segmentation given ground truth. While [10] utilizes the number of missclassified pixels, this work utilizes curve metrics. The Sobolev distance [11] is a shape metric for curves implicitly defined by a signed distance function; it computes point-wise errors between the two curves' signed distances. The Sobolev distance provides a local measure of curve mismatch. The Laplace distance [14] is a metric on the space of curves that locally provides the distance between curves, by computing the length of unique correspondence trajectories between the two curves.

3) Segmentation error vs contrast parameter: Here, we study the influence of noise on the segmentation process and use the Bhattacharyya distance in order to predict the segmentation error. Each sensor and visual tracking application will have different noise level characteristics and tolerances. The process presented here should be viewed as an important calibration step to perform before using a segmentation algorithm for tracking: it characterizes the nominal performance and response to imaging noise. The empirical uncertainty calibration is described in what follows, where we use Bayesian segmentation [4], [5] on the image data.

Protocol: First, begin with a collection of shapes that will form the ground truth (preferably from an existing video sequence). Select the interior and exterior distributions to be Gaussian, $P_{in,out} = \mathcal{N}(\cdot; \mu_{in,out}; \sigma_{in,out})$. with μ_{in}/μ_{out} and σ_{in}/σ_{out} the interior/exterior Gaussian parameters. Add zero mean Gaussian noise with standard deviation $\sigma_{noise} > 0$ to the images. For each choice of σ_{noise}, generate a set of corrupted images. Perform segmentation to yield curves partitioning the images into target and background regions. Determine the contrast coefficient, as given by the Bhattacharya distance between the interior and exterior distributions of pixel intensities, using ground truth. Compute the curve estimation error using the Sobolev and Laplace shape metrics. With these measurements, derive the expected segmentation error as a function of the Bhattacharyya distance.

Experiment: We used a collection of 36 different shapes both artificially generated, and hand-segmented from real images. The collection of shapes considered included circles of different radii, walking people, and fishes. For each noise level, σ_{noise}, 180 realizations of noisy images were generated. Figure 2 depicts the experimental curves giving the segmentation error as a function of the contrast parameter. The mean curve is given by the thick line curve, and serves as a first approximation to the segmentation error given the Bhattacharyya coefficient. Using the fit, the Bhattachrya measure will map to the expected segmentation error.

When the target and background are clearly separable ($\beta < 0.4$), the error dependence on β is independent of shape. Due to the clear separation, the segmentations have low error. A Bhattacharyya coefficient between 0.4 and 0.7 represents the transition region from moderate to poor separation of target and background. In this range,

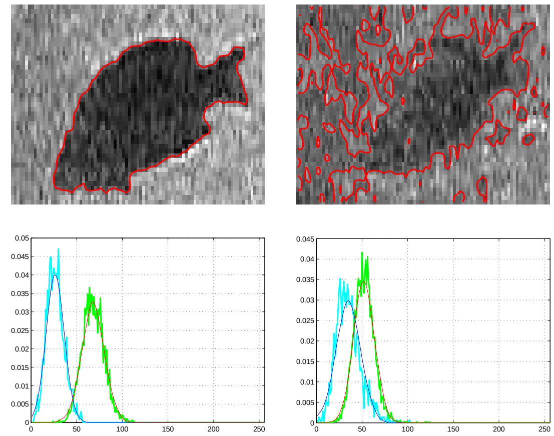


Fig. 1. Image samples and corresponding taget/background distributions. The left column represents a scenario where the pdf's are clearly separated ($\beta = 0.12$). The right column represents a scenario where the pdf's overlap significantly ($\beta = 0.75$). The true pdf's are given by thick lines while Gaussian-fitted pdf's are shown with fine lines.

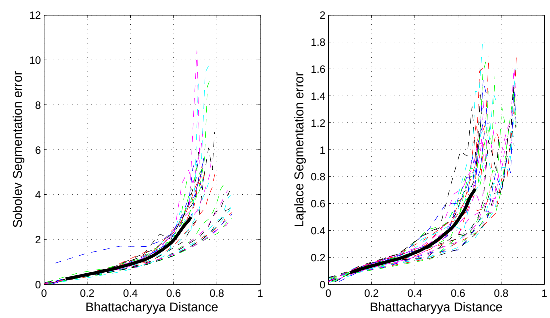


Fig. 2. Segmentation error as a function of the Bhattacharyya coefficient between the target and background distributions

the spread between the curves is larger. Still, the mean curve provides a consistent measure of the expected segmentation error given the Bhattacharya measure. The wider spread is due to an increased dependency on the shape. For sufficiently low noise levels, the segmentation error is fairly independent of target shape. For significant noise levels, the segmentation algorithm performance also depends on the local curvature of the shape (be it high or low). Such behavior is expected given that many segmentation methods utilize curve smoothing priors during the optimization process. The error spread for high β values reflects the dependence of the error on the local shape curvature and the influence of the curve smoothing terms in the segmentation algorithm. Above a certain noise level ($\beta > 0.7$), target and background are no longer separable; the segmentation results are meaningless (the smoothing terms dominate).

III. An Adaptive Probabilistic Filter

This section describes the probabilistic filtering strategy used to spatio-temporally constrain the individual segmentations from a video sequence. Essential to many filtering strategies is a gain factor that modulates the influence of the measurement versus the prediction [12]. The error characterization process described above assessed the measurement uncertainty asociated to imaging noise. A mapping of the measurement uncertainty to the optimal gain is now sought. In what follows, we describe the probabilistic filtering strategy and the empirical evaluation of the optimal gain. It is a modification of the deformotion filter [7], and is described further in [13].

4) Filtering Strategy: The visual track filter operates via interlaced prediction and correction steps. The filter decomposes the tracking procedure into two components: a rigid state (pose) and a non-rigid state (shape). The pose consists of translation, g and translational velocity, ξ, in $\mathbb{R}^2$. The shape state implicitly represents the curve by a probability field, P. The probability field implicitly defines the shape curve by the 50% probability contour. The shape velocity is a vector field on $\mathbb{R}^2$, Θ.

The prediction step evolves the observer state using a constant velocity model:

$$\begin{aligned} \dot{g} &= \xi, & \dot{\xi} &= 0, \\ \dot{P} &= \nabla P \cdot \Theta, & \dot{\Theta} &= 0, \end{aligned}$$

to generate the predicted states $(g^-, \xi^-, P^-, \Theta^-)$.

Given a segmentation of the current image, the measurement for the current image is generated as follows. Align the predicted shape with the segmented shape, which will provide the translational component, g_m, of the measured shape. The segmentation, after registration and possibly also conversion to implicit density form, provides the shape measurement, P_m. If desired, measurement of the shape velocities, Θ_m, may be obtained through the use of optical flow [6] between the current and previous images.

Once the measurement is complete, a correction procedure is applied to generate the current estimate of the track signal, with decoupled pose and shape corrections. The pose correction is performed using a Kalman filter update on the combined pose and pose velocity (g, ξ) sub-state. The shape correction is done using a geometric filtering strategy on the probability field and its complement, Q,

$$\begin{aligned} P^+ &= (P^-)^{1-K} \cdot (P_m)^K \\ Q^+ &= (1 - P^-)^{1-K} \cdot (1 - P_m)^K, \end{aligned}$$

after which P^+ and Q^+ are re-normalized, but only P^+ is kept. The parameter K varies in the range $[0, 1]$; lower values favor prediction and higher values favor measurement.

Correction on the velocity field is given by

$$\Theta^+ = \Theta^- + K_{vx} \cdot X_{err}(P_m, P^-) + K_{vv} \cdot (\Theta_m - \Theta^-).$$

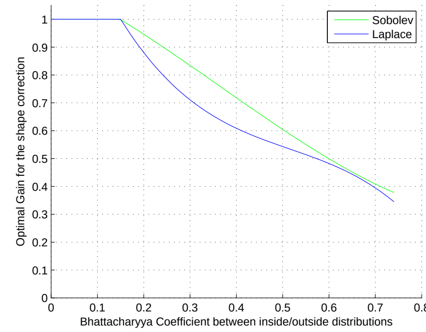


Fig. 3. Optimal gain as a function of the Bhattacharyya coefficient.

One way to generate the error vector field between probability fields $X_{err}(P_m, P^-)$ is to compute the optical flow between the measured and predicted shapes. The parameters K_{vx} and K_{vv} vary in $[0, 1]$. If the shape velocities are not measured, then $K_{vv} = 0$.

Due to the nonlinear nature of shape, there may be a nonlinear relationship between the expected segmentation error (via Sobolev of Laplace metrics) and the optimal shape correction gain, K, during tracking. A second experiment was performed to identify the relationship between the Bhattacharya coefficient and the optimal gain. Because the gains K_{vx} and K_{vv} operate in a vector space, we do not optimize them.

Protocol: Take video sequences with ground truth and inject a known amount of noise, σ_{noise} into the sequences similar to before. For gain values K in the range $[0, 1]$, perform the experiment at each realization of the noise level (with fixed K_{vx} and K_{vv}). Quantify, via the Sobolev or Laplace metrics, the tracking performance of the filter at the different gain levels. Collect the results together to obtain the optimal gain as a function of the Bhattacharya measure.

Experiment: The protocol was followed for a single image sequence. There were 240 different configurations for the inside/outside distributions, and the gain sweep went in increments of 0.05. The resulting functional dependence is given in Figure 3. The optimal gain for the Sobolev metric gives an almost linear dependence for about $\beta > 0.15$. The optimal gain for the Laplace metric has slight nonlinear dependence, but approximately follows the trend of the Sobolev metric optimal gain.

IV. Experiments and Results.

For the experiments, we used a synthetically corrupted infrared sequence from OTCBVS and a naturally noisy aquarium sequence. The sequences were tracked with constant filter gains. Then we used the Bhattacharyya distance to adjust the gain. Because it is unrealistic to assume ground truth is available, we computed the Bhattacharyya distance with the interior/exterior distributions generated by the segmentations. Performance evaluation used the Laplace metric in conjunction with hand-segmentations (ground truth) of the sequences.

Figure 4 shows the results obtained on the infrared sequences for different noise levels. The adaptive gain has good overall performance compared to fixed gains. For naturally noisy sequences with variable noise levels as in the aquarium sequence (frames 10 and 60 depicted in Figure 1), the Bhattacharyya coefficient proves to be efficient at triggering shape correction when necessary and assessing the extent to which such correction needs to be performed. Figure 5 shows the Laplace error as a function of time. The adaptive filter tracks well throughout.

V. Conclusion

This paper presented a procedure to characterize the behavior of segmentation algorithms in the presence of noise. The Bhattacharyya coefficient between target and background distributions proved to be useful for assessing segmentation error. Experiments on noisy sequences verified the ability of the Bhattacharyya distance to assess imaging noise for adaptive shape filtering. Future work seeks to perform the same analysis for color images.

References

[1] G. Charpiat, O. Faugeras, and R. Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Found. of Comp. Math.*, 5(2/2):1–58, 2004.

[2] D. Cremers. Dynamical statistical shape priors for level set-based tracking. *IEEE TPAMI*, 28:1262–1273, 2006.

[3] S. Dambreville, Y. Rathi, and Tannenbaum A. Tracking deformable objects with unscented Kalman filtering and geometric active contours. In *ACC*, pages 2856–2861, 2006.

[4] S. Haker, G. Sapiro, and A. Tannenbaum. Knowledge-based segmentation of SAR images. In *ICIP*, pages 597–601, 1998.

[5] S. Haker, G. Sapiro, A. Tannenbaum, and D. Washburn. Missile tracking using knowledge-based adaptive thresholding: Tracking of high speed projectiles. In *ICIP*, pages 786–789, 2001.

[6] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[7] J.D. Jackson, A.J. Yezzi, and S. Soatto. Tracking deformable moving objects under severe occlusions. In *CDC*, pages 2990–2995, 2004.

[8] F. Lecellier, S. Jehan-Besson, J. Fadili, G. Aubert, M. Revenu, and E. Saloux. Region-based active contour with noise and shape priors. In *ICIP*, pages 1649–1652, 2006.

[9] T.C.M. Lee. Segmenting images corrupted by correlated noise. In *ICIP*, pages 247–250, 1997.

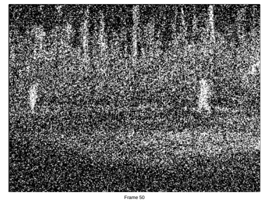
[10] P. Matin, P. Refregier, F. Goudail, and F. Guerault. Influence of the noise model on level set active contour segmentation. *IEEE TPAMI*, 26(6):799–803, 2004.

[11] P. Maurel, R. Keriven, and O. Faugeras. Reconciling landmarks and level sets. In *ICPR*, volume 4, pages 69–72, 2006.

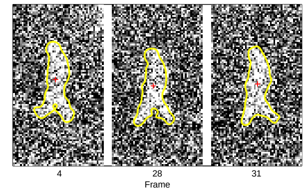
[12] P. S. Maybeck. *Stochastic models, estimation and control. Volume I.* Press, Academic, 1979.

[13] I.J. Ndiour, O. Arif, J. Teizer, and P. Vela. A probabilistic shape filter for online contour tracking. *accepted in ICIP*, 2009.

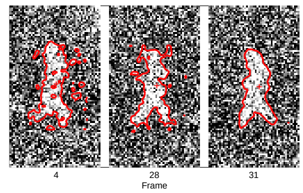
[14] E. Pichon, D. Nain, and M. Niethammer. A Laplace equation approach for shape comparison. *SPIE Medical Imaging*, 2:24–132, 2006.



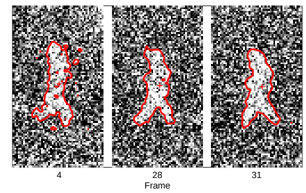
(a) Sample frame (320x240)



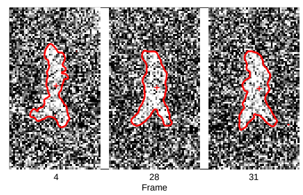
(b) Ground Truth



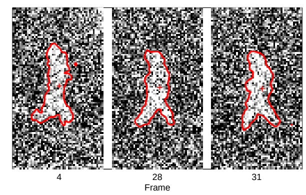
(c) Fixed Gain K = 1



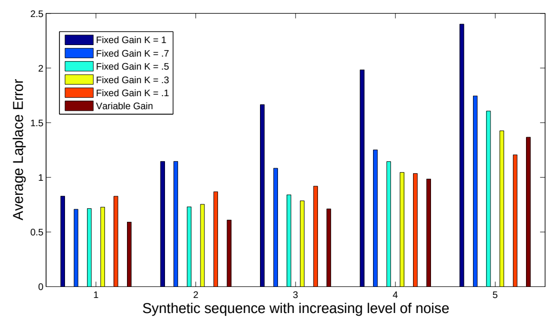
(d) Fixed Gain K = 0.7



(e) Fixed Gain K = 0.4

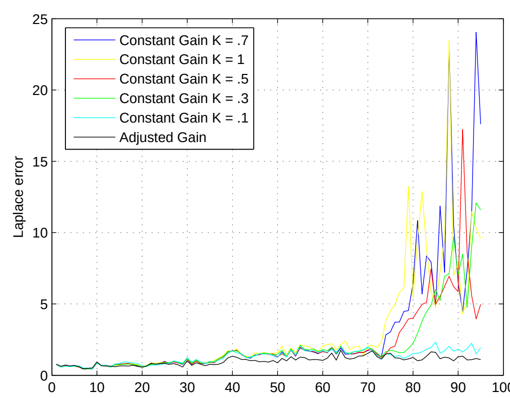


(f) Variable Gain



(g) Laplace error

Fig. 4. For one level of noise corruption, image sample and segmentations obtained at given times for different values of the gain. For five levels of noise corruption, the average Laplace errors throughout the sequences are displayed.



(a) Laplace error vs. time

Fig. 5. Noisy aquarium sequence.