Tracking with a New Distribution Metric in a Particle Filtering Framework

Romeil Sandhu[a] Tryphon Georgiou[b] Allen Tannenbaum[a]

[a]School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250

[b]Department of Electrical & Computer Engineering, University of Minnesota, Minneapolis, MN 55455

ABSTRACT

Tracking involves estimating not only the global motion but also local perturbations or deformations corresponding to a specified object of interest. From this, motion can be decoupled into a finite dimensional state space (the global motion) and the more interesting infinite dimensional state space (deformations). Recently, the incorporation of the particle filter with geometric active contours which use first and second moments has shown robust tracking results. By generalizing the statistical inference to entire probability distributions, we introduce a new distribution metric for tracking that is naturally able to better model the target. Also, due to the multiple hypothesis nature of particle filtering, it can be readily seen that if the background resembles the foreground, then one might lose track. Even though this can be described as a finite dimensional problem where global motion can be modeled and learned online through a filtering process, we approach this task by incorporating a separate energy term in the deformable model that penalizes large centroid displacements. Robust results are obtained and demonstrated on several surveillance sequences.

Keywords: Geometric Active Contours, Distributions, Particle Filtering, Tracking, Metrics

1. INTRODUCTION

A well-studied problem in computer vision is the fundamental task of tracking moving and deformable objects.[1,2] In this paper, we adopt the idea of motion described by Yezzi and Soatto.[3] We decouple motion into two distinct parts: a "global" rigid motion, and a "deformation" which is given by any departure from rigidity. However, before detailing the approach taken, we briefly revisit similar approaches and schemes of previous related work. Also, it should be noted that the main contribution in this paper is the introduction of a new distribution metric for visual tracking. We refer the interested reader to a rigorous development of the metric[4,5]

The first class of tracking schemes propose the use of a finite dimensional representation of continuous curves. Specifically, the B-Spline representation is often used for the "snake" model.[2] Isard and Blake apply this representation for the contour of the objects and proposed the Condensation algorithm.[1,6] By assuming a unimodal distribution of the state vector, the unscented Kalman Filter is similarly introduced in combination with B-Splines for rigid object tracking.[7,8] Although these approaches track the finite dimensional group (e.g., Euclidean, affine) parameters, they do not explicitly handle the local deformation of an object.

The use of an implicit representation of the contour via level set methods[9,10] comprise a second class of algorithms. Given an initial estimate, a curve is evolved until it minimizes an image based energy functional. Several energy functionals, which utilize different features of an image, can be found in literature.[11–14] Moreover, level set methods used primarily for tracking exploit both the finite dimensional group and local deformations.[3,15–18] Specifically, Freedman *et. al.*[19] propose to track with distributions. Given a prior knowledge of the object, the algorithm seeks to match distributions via the Bhattacharyya measure or Kullback-Leibler divergence. However, whereas these distance functionals originate from Information Geometry, our similarity metric is based on prediction theory.

Further author information: (Send correspondence to R.S.)
R.S.: E-mail: rsandhu@gatech.edu, Telephone: 1 404 385 5062

Image Processing: Machine Vision Applications, edited by Kurt S. Niel, David Fofi,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6813, 68130N, © 2008 SPIE-IS&T · 0277-786X/08/$18

Recently, Rathi *et. al.*[17,20] use particle filtering in conjunction with active contours that utilize first and second moments for tracking deformable objects. The algorithm is further extended to incorporate shape information.[21] To reduce the computational complexity involved with the particle filtering scheme, Dambreville *et. al.*[22] propose an unscented Kalman Filtering approach. However, the authors note that the algorithm suffers the limitation of assuming a unimodal probability distribution of the state vector, and thus, may fail for multimodal distributions.

In this paper, we introduce a new distribution metric for tracking that is naturally able to better model the target or object of interest. This intrinsic metric quantifies "distance" for two density functions as the standard deviation of the difference between logarithms of those density functions. Also, because of the temporal nature of surveillance data, filtering theory becomes increasingly relevant. Using the particle filtering framework proposed by Rathi *et. al.*,[17] we incorporate this density metric as apart of the deformable model through the use of Geometric Active Contours (GAC). Moreover, a separate energy term that penalizes large centroid displacements is added to this deformable model. In doing so, we are able to handle the multiple hypothesis nature of particle filtering by biasing particles in region near the predicted state. That is, large displacements for the centroid are weighted against the probability of how well it matches our known target. Finally, it should be noted that the dynamical model for the finite dimensional group is simplified in order to highlight the influence of tracking with the newly proposed metric.

This paper is organized as follows: In the next section, we derive a variational framework for matching two densities using the proposed distribution functional as well as a centroid penalization scheme. This is followed with a brief review of particle filtering theory. In Section 3, we describe the tracking algorithm along with the specifics of the state space, prediction model, measurement model, and resampling scheme. Experimental results are given in Section 4. We conclude with a brief summary in Section 5.

2. PRELIMINARIES

In this section, we derive a variational framework for our proposed metric and for centroid penalization as well as review some basic notions from the theory of particle filtering, which we will need in the sequel.

2.1 A Similarity Metric for Region Based Segmentation

We begin by casting the proposed distribution functional in the GAC framework. However, before doing so, we briefly motivate our reason for such a distribution metric as well as provide an overview of theory behind its resulting similarity measure.

2.1.1 Motivation and Theoretical Overview

As stated previously, we seek to match a distribution of a certain photometric variable (e.g., gray-scale intensity, color, tensor) with a known distribution. In other words, we look to "predict" the optimal density that best correlates with a given distribution. Originally motivated by measuring the similarity between spectral distributions, the metric is derived based on principals in prediction theory. That is, let $f_1(\theta)$ and $f_2(\theta)$ denote the power spectral distribution defined on the interval $\theta \in [0, \pi]$. These distributions correspond to their respective zero-mean random stationary processes $u_k^{f_i}$ with $i \in \{1, 2\}$ and $k \in \mathbb{Z}$. Then the variance of the linear or one-step-ahead prediction error for a given process of spectral distribution $f_i(\theta)$ is

$$\mathcal{E}\{|u_0^{f_i} - \hat{u}_{0|past}^{f_i}|^2\} = \mathcal{E}\{|u_0^{f_i} - \sum \alpha_k^{f_i} u_{-k}^{f_i}|^2\} \tag{1}$$

with $k > 0$ and where $\alpha_k^{f_i}$ are the coefficients that minimize the linear prediction error variance for the specific $f_i(\theta)$. Now suppose that we are given a known power distribution $f_1(\theta)$, and we would want to measure the "distance" or similarity with another spectral density $f_2(\theta)$. This can be achieved by first assuming the density $f_2(\theta)$ originates from the distribution $f_1(\theta)$. From this, we use $f_2(\theta)$ to design a predictor and compare how well it performs **against the optimal prediction** that is based on $f_1(\theta)$. Hence, we arrive at the following measure, which is denoted as the *degradation of predictive error variance* and is given below

$$\rho(f_1, f_2) = \frac{\mathcal{E}\{|u_0^{f_1} - \sum_{k=1}^{k=\infty} \alpha_k^{f_2} u_{-k}^{f_1}|^2\}}{\mathcal{E}\{|u_0^{f_1} - \sum_{k=1}^{k=\infty} \alpha_k^{f_i} u_{-k}^{f_1}|^2\}} \tag{2}$$

Equation (2) gives us a basis for measuring (dis)similarity and can be viewed as analogous to "divergences" such as the Kullback-Leibler entropy seen in Information Theory. Moreover, in the particular context of visual tracking, our basis for seeking a matching distribution given a known distribution arises naturally from the principal motivation of prediction as discussed above. As a result of the above formulation, Equation (3) is the derived metric given in the level set framework. While the derivation and proof of the metric is beyond the scope of this paper, we do refer the interested reader for an enlightened discussion on its properties, derivation, and the parallels to Information Geometry, including both the Bhattacharyya measure and Kullback-Leibler divergence.[4,5]

2.1.2 Level Set Formulation

We consider the problem of segmenting an image I. That is, we first assume the image is composed of two homogeneous regions referred to as "Object" and "Background". From this, the goal of segmentation is to capture these two regions. To do so, we enclose a curve $\mathcal{C}$, represented as the zero-level set of a signed distance function $\phi : \Re^2 \rightarrow \Re$, such that $\phi < 0$ represents the inside of $\mathcal{C}$ and $\phi > 0$ represents the outside of $\mathcal{C}$. Our goal is to evolve the curve $\mathcal{C}$, or equivalently ϕ, so that the interior matches the Object given by the distribution q_{known}. The curve $\mathcal{C}$ would then match the boundary $\partial\Omega$ separating the Object and Background.

We now propose an energy functional based on the metric discussed previously,[5] and derive the corresponding partial differential equation (PDE) that describes its curve evolution in the level set framework. It should be noted that segmentation results are dependent on the distribution metric used. Given prior (or learnt) knowledge of the distribution of the object q_{known}, we seek to measure the similarity between distributions through the following energy functional

$$T(z,\phi) = \sqrt{\int_z \left[z\log\frac{p_{object}(z,\phi)}{q_{known}(z)} \right]^2 dz - \left[\int_z z\log\frac{p_{object}(z,\phi)}{q_{known}(z)} dz \right]^2} \tag{3}$$

where $z \in \mathcal{Z}$ is the photometric variable, and q_{known} and p_{object} are the probability density function's (pdf) defined on the random variable z. In the present work, we restrict the variable z to set of gray level values $\{1, 2, ..., 256\}$. Moreover, let $I : \Re^2 \rightarrow \mathcal{Z}$ be a mapping of the image defined over the domain Ω, to the photometric variable z, and $x \in \Re^2$ be the image coordinates. Then the pdf inside the curve $\mathcal{C}$ can be formulated as such

$$p_{object}(z,\phi) = \int_\Omega \frac{K(z - I(x))H(-\phi)}{H(-\phi)} dx \tag{4}$$

where $K(z - I(x))$ is a specified kernel. For numerical experiments, we have used $K(z - I(x)) = \delta(z - I(x))$. Also $H_\epsilon : \mathbf{R} \mapsto \{0, 1\}$ denotes the Heaviside step function with the corresponding derivative δ_ϵ. These are both given as follows

$$H_\varepsilon(\phi) = \begin{cases} 1 & \phi < \epsilon \\ 0 & \phi > \epsilon \\ \frac{1}{2}(1 + \frac{\phi}{\epsilon} + \frac{1}{\pi}\sin(\frac{\pi\phi}{\epsilon})) & \text{otherwise} \end{cases} \qquad \delta_\epsilon(\phi) = \begin{cases} 0 & \phi < \epsilon, \quad \phi > \epsilon \\ \frac{1}{2\epsilon}(1 + \cos(\frac{\pi\phi}{\epsilon})) & \text{otherwise} \end{cases} \tag{5}$$

The gradient $\nabla_\phi T$ can be computed using the calculus of variations. Taking the first variation with respect to ϕ yields the following PDE

$$\nabla_\phi T = -\frac{\delta_\epsilon(\phi)}{T} \cdot [\mathcal{E}\{B \cdot G\} - \mathcal{E}\{B\} \cdot \mathcal{E}\{G\}] \tag{6}$$

with B and G given as

$$B = \log\frac{p_{object}(z,\phi)}{q_{known}(z)} \qquad G = \frac{\delta(\phi)}{A_{in}}\left(1 - \frac{1}{p_{object}(z,\phi)}\right)$$

Noting that A_{in} is given by $\int_\Omega H_\epsilon(\phi)dx$, Equation (6) is a PDE that describes the evolution of the curve $\mathcal{C}$ that optimally minimizes the "distance" between an *a-priori* distribution with that of the pdf inside the curve. Next, we formulate a centroid penalization variationally in the similar manner as above.

2.2 A Variational Penalization Scheme

Due to the multiple hypothesis nature of particle filtering, if the background resembles the modeled distribution q_{known}, one might lose track (i.e. track another object apart of the background). This is caused by the definition of the importance weights and the selection of the measurement, which is defined by the image energy. In this section, we introduce an energy functional that penalizes the centroid with respect to a known position. The motivation for such an energy is to bias the evolution of the curve in Equation (6), resulting in an overall flow that is not purely dependent on the photometric variable z. Moreover, in a particle filtering scheme, this penalty and the number of R iterations performed (see Section 3.4) will be a parameter that can represent how much one trusts the system model versus the obtained measurement.

Again, let $x \in \Re^2$ be the coordinates of image I defined over the domain Ω. Also let $\hat{c}_{object}$ denote the centroid of the object for which we would like to penalize large deviations. Defining the centroid of the curve as $c_{curve}(x,\phi) = \int_\Omega \frac{xI(x)H(-\phi)}{H(-\phi)}dx$, we formulate an energy functional as follows

$$B(x,\phi) = \sqrt{\hat{c}_{object}(x) - c_{curve}(x,\phi)} \tag{7}$$

Taking the first variation of Equation (7), yields the following minimization

$$\nabla_\phi B = -\frac{\delta_\epsilon}{B}\left[(\hat{c}_{object} - c_{curve})\left(\frac{c_{curve} - xI(x)}{A_{in}}\right)\right] \tag{8}$$

Combining Equation (6) and Equation (8) with a regularizing term commonly used, we arrive at the following PDE

$$\nabla_\phi E_{image} = -\delta_\epsilon(\phi)\left[\nabla_\phi T + \lambda_{centroid}\cdot\nabla_\phi B - \lambda_{smooth}\cdot \mathrm{div}\left(\frac{\phi}{|\phi|}\right)\right] \tag{9}$$

The weighting parameters $\lambda_{centroid}$ and λ_{smooth} are chosen dependent on method of segmentation as well as the type of imagery. In this paper, they are chosen subjectively in the range [0,1]. Next, we briefly review basic theory from particle filtering.

2.3 Particle Filtering

Letting $x \in \mathbb{R}^n$, Monte Carlo methods allow for the evaluation of a multidimensional integral $I = \int g(x)dx$ via a factorization of the form $I = \int f(x)\pi(x)dx$, whereby $\pi(x)$ can be interpreted as a probability distribution. Taking samples from such a distribution in the limit yields the estimate of I that would otherwise be difficult or impossible to compute. However, generating samples from the posterior distribution is usually not possible. Thus, if one can only generate samples from a similar density $q(x)$, the problem becomes one of "importance sampling." That is, the Monte Carlo estimate of I can be computed by generating $N >> 1$ independent samples $\{x^i; i = 1, ..., N\}$ distributed according to $q(x)$ by forming the weighted sum: $I_N = \frac{1}{N}\sum_{i=1}^N f(x^i)w(x^i)$, where $w(x^i) = \frac{\pi(x^i)}{q(x^i)}$, represents the normalized importance weight. Thus, by employing Monte Carlo methods in conjunction with Bayesian filtering, authors[23] first introduced the Particle Filter. We refer the reader for an in-depth discussion and popular texts on Monte Carlo methods and particle filtering schemes.[24,25]

Now considering $x_t \in \mathbb{R}^n$ to be a state vector, particle filtering is a technique for implementing a recursive Bayesian filter through Monte Carlo simulations. At each time t, a cloud of N particles is produced, $\{x_t^i\}_{i=1}^N$, whose empirical measure closely "follows" $p(x_t|z_{0:t}) = \pi_t(x_t|z_{0:t})$, the posterior distribution of the state given the past observations.

The algorithm starts with sampling N times from the initial state distribution $\pi_0(x_0)$ in order to approximate it by $\pi_0^N(x_0) = \frac{1}{N}\sum_{i=1}^N \delta(x_0 - x_0^i)$, and then implements Bayesian recursion at each step. With the above formulation, the distribution of the state at $t-1$ is given by $\pi_{t-1}(x_{t-1}|z_{0:t-1}) \approx \frac{1}{N}\sum_{i=1}^N \delta(x_{t-1} - x_{t-1}^i)$. The algorithm then proceeds with a **prediction step** that draws N particles from the proposal density $q(x_t|z_{0:t-1})$. With appropriate importance weights assigned to each particle, the *prediction distribution* can now be formed in a similar fashion as above, i.e. $\hat{\pi}_t(x_t|z_{0:t-1}) = \frac{1}{N}\sum_{i=1}^N w_t^i \delta(\hat{x}_t - \hat{x}_t^i)$. Then, in the **update step**, new information

arriving online at time t from the observation z_t is incorporated through the importance weights in the following manner:

$$w_t^i \propto w_{t-1}^i \frac{p(z_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{t-1}^i, z_t)}. \tag{10}$$

From the above weight update scheme, the *filtering distribution* is given by $\tilde{\pi}_t(x_t|z_{0:t}) = \frac{1}{N}\sum_{i=1}^{N} w_t^i \delta(x_t - x_t^i)$. Resampling N times with replacement from $\tilde{\pi}_t$ allows us to generate an empirical estimate of the posterior distribution π_t. Even though $\tilde{\pi}_t$ and π_t both approximate the posterior, resampling helps increase the sampling efficiency as particles with low weights are generally eliminated.

3. TRACKING ALGORITHM

In this section, we incorporate the derived flow in Equation (9) into a similar particle filtering framework proposed by Rathi *et. al.* Specifics are given on the state space model, prediction model, measurement model and the resampling scheme used.

3.1 State Space Model

As stated in the prequel, we are interested in tracking two types of motion: a "global" rigid motion, and "deformation" which is any departure from rigidity. Moreover, Yezzi and Soatto show that the overall motion can be described by a set of non-unique rigid motion parameters along with a "deformation" function. From this, we assume that the "global" motion of an object is given by the translation of its centroid, and any other "deformation" is captured by the curve evolution in Equation (9) described in the previous section. Justification for doing curve evolution can be found in literature, specifically particle filtering.[17,20] The state vector can now be defined in the following manner

$$x(t) = \begin{pmatrix} x_c \\ y_c \\ \mathcal{C} \end{pmatrix}(t) = \begin{pmatrix} \gamma \\ \mathcal{C} \end{pmatrix}(t) \tag{11}$$

where $\gamma = [x_c, y_c]^T$ is the object's centroid, and $\mathcal{C}$ denotes the contour of deformation represented by the zero level set of ϕ. After forming an estimate through the prediction model from the distribution $p(x_t|x_{t-1}, z_{t-1})$, we obtain an observation $z(t) = I(t)$ which is the image received online at time t. Thus, the observation space is given as follows:

$$z(t) = \begin{pmatrix} x_c^m \\ y_c^m \\ \mathcal{C}^m \end{pmatrix}(t) = \begin{pmatrix} \gamma^m \\ \mathcal{C}^m \end{pmatrix}(t) \tag{12}$$

where γ^m and $\mathcal{C}^m$ are the measured centroid and deformation, respectively.

3.2 Prediction Model

For a given set of particles $\{x^i; i = 1, ..., N\}$, we seek a model for the prediction distribution, which predicts both the "global" rigid motion and the "deformation" of an object. In this paper, two departures taken from the general particle filtering framework (in addition to the contributions mentioned in this paper) are to simplify the dynamical model for predicting rigid motion as well as assume a translational prior for the proposal distribution. Note, this is only done to highlight the influence of the new distribution metric for visual tracking. From this, the proposal distribution is given as follows:

$$\begin{aligned} q(x_t|x_{t-1}^i, z_k) &= p(x_t|x_{t-1}^i) \\ &= p(\gamma_t|\gamma_{t-1}^i)p(\mathcal{C}_t|\mathcal{C}_{t-1}^i) \end{aligned} \tag{13}$$

In particular, $\hat{\gamma}_t^i = \gamma_{t-1}^i + v_{t-1}^{\gamma}$, where v_{t-1}^{γ} is zero-mean gaussian random noise that is added to the spatial location of the centroid and the predicted contour of deformation is the prior contour given at time $t-1$, $\hat{\mathcal{C}}_t^i = \mathcal{C}_{t-1}^i + v_{t-1}^{\mathcal{C}}$. In this paper, we take $v_{t-1}^{\mathcal{C}} = 0$, causing the prediction of the contour to be deterministic.

Although adding noise to an infinite dimensional representation of a contour is generally not easy, one can perform Principal Component Analysis (PCA) and add noise in the principal variation directions. One can also incorporate a shape prior in the prediction step. The simplification of the weight update scheme is now made by substituting Equation (13) into Equation (10) as

$$w_t^i \propto w_{t-1}^i p(z_t|\hat{x}_t). \tag{14}$$

Given the estimate $\hat{x}_t$, we now proceed in describing the measurement model, which includes the curve evolution of Equation (9).

3.3 Measurement Model

The measurement function, $z_t = h\big(\hat{x}_t, I(t)\big)$, where $\hat{x}_t$ is a seed point (corresponding to a curve centered at a certain position), and $I(t)$ is the image that becomes available at time t, can be described as follows:

1. Run minimization of the functional (9) for R iterations for each of the $\hat{x}_t^i$: the choice of R depends on the type of metric used or energy being minimized. This results in a local exploration of both the position and deformation. Also note, $\hat{c}_{object}$ in functional (9) is the measured centroid at time t.
2. Compute an update of the importance weight by Equation (14) by defining $p(z_t|\hat{x}_t) \equiv e^{-E_{image}(\mathcal{C}_t, z_t)}$.
3. Build a cumulative distribution function from these importance weights. Using the generic method, resample N times with replacement to generate N new samples.[24]
4. Select the curve with the minimum energy (best fitting curve) as the measurement. That is, the centroid of the selected curve as well as the deformation of that curve are taken as the measurement for the system.

As can be noticed from above, the posterior distribution can be multi-modal and the measurement function is highly non-linear because of curve evolutions and selection of the best fitting curve. Thereby, this justifies the use of a particle filter to track as compared to unscented Kalman Filtering or the more classical Kalman Filtering. Next, we discuss the resampling scheme, and the importance of doing gradient descent for R iterations.

3.4 Resampling

The resampling step is introduced into particle filtering schemes as a solution to "sampling degeneracy," which is unavoidable in sequential importance sampling. That is, it can be shown show that the variance of importance weights are only allowed to increase over time.[26] This results in particles that are not concentrated in a region of high likelihood of the posterior distribution. Aside from the computational cost, the phenomenon of degeneracy creates poor filtering results. Thus, in this paper, we adopt the general resampling scheme.[24]

Although resampling attempts to solve "sampling degeneracy," it induces another problem known as "sample impoverishment," whereby all of the particles are *only* concentrated in a single region. Because this leads to a loss of diversity for a given set of particles, an approximation to the posterior distribution is not accurate, and the registration may fail.

To address both of the above problems, careful consideration of the number of curve evolutions R must be made. By choosing R too large, we would be converging towards the local minima. This is not desirable since the state at t and $t-1$ would lose dependency. Indeed, "sample degeneracy" will occur if all of the particles tend toward one region. Likewise, if R is chosen to be too small, then most of the particles may never be associated with the high likelihood region of the posterior resulting in "sample impoverishment." In other words, *the choice of R depends on how much one trusts the system model versus the obtained measurement.* For our experiments, we have found that a range of $R = [2, 5]$ has given robust results, as will be seen in Section 4.

4. EXPERIMENTS

In this section, we present experimental results of the proposed algorithm on varying surveillance sequences. In particular, we demonstrate the importance of generalizing the statistical inference by tracking a person with a multimodal distribution through a crowded sequence. Moreover, we provide an example in which the multiple hypothesis nature of particle filtering becomes ill-suited in a clutter environment. This is corrected via the centroid penalization scheme.

4.1 Cluttered Environments

In this example, we demonstrate the effectiveness of using the proposed centroid penalization scheme within the deformable model of the particle filtering framework. The goal here is to continuously track a car from an aerial sequence through a clutter environment, which share similar statistical information to that of the object being tracked. Assuming simplistic dynamics for predicting the global motion, Figure 1 shows that if the background resembles the foreground, we lose track. This is caused by the selection of the measurement via the image energy and importance weights. By defining weights that are dependent on only statistical information, tracking through clutter becomes difficult if no special consideration is made.

Introducing the centroid penalization variationally within the deformable model not only allows the prediction step to be driven by a geometric constraint, but affects the update of the importance weights. With the same conditions as applied in Figure 1, we successfully track the car through clutter as seen in Figure 2. However, we note that an important advantage of using particle filtering with GAC is to capture erratic motion and the parameter $\lambda_{centroid}$ must be adjusted appropriately depending on the type of imagery. We chose $R = 2$ with 40 particles.

4.2 Multimodal Crowd Sequence

While a Gaussian assumption can generally be made on an object of interest, it may not hold for a multimodal target. That is, Figure 3 shows a person walking through a crowded sequence. In particular, as the object cross over the darker gray portion, the contour may leak and not accurately capture the person. By learning the distribution online and prior to a given time t, we are able to accurately quantify the foreground to that of the background. In doing so, we use 100 particles with $R = 3$. We show several screen shots from the surveillance sequence as the person walks through a courtyard.

4.3 Camera Induced Deformation

In this example, we show the robustness of the algorithm by tracking (with a simplified dynamical model) when there exists camera motion. Here, we treat this motion as a "deformation" occurring to the object itself. As stated previously, an advantage of particle filtering with GAC is the ability to capture erratic motions. The parameter $\lambda_{centroid}$ is reduced and v_t^{γ}, the zero-mean gaussian noise applied to the centroid, is increased. Hence, we must rely on the discriminating power of the distribution metric. Figure 4 shows several screen shots taken that tracking successfully through camera motion.

5. CONCLUSION

In this work, our contributions are twofold. First we extend the deformable model of the filtering framework by introducing our distribution metric, which makes its first appearance in computer vision for visual tracking. The metric quantifies "distance" for two density functions as the standard deviation of the difference between logarithms of those density functions. Assuming that the target's probability distribution is known a priori, we proposed to minimize the distance of this known distribution with a region enclosed by our contour via gradient descent. Secondly, we added a variational penalization scheme on the centroid. Thus, large displacements for the centroid are weighted against the probability of how well it matches our known target. This tradeoff is important and is adjusted appropriately in achieving robust results on varying surveillance sequences.

ACKNOWLEDGMENTS

This work was supported in part by grants from NSF, AFOSR, ARO, MURI, as well as by a grant from NIH (NAC P41 RR-13218) through Brigham and Women's Hospital. This work is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149. Information on the National Centers for Biomedical Computing can be obtained from `http://nihroadmap.nih.gov/bioinformatics`.

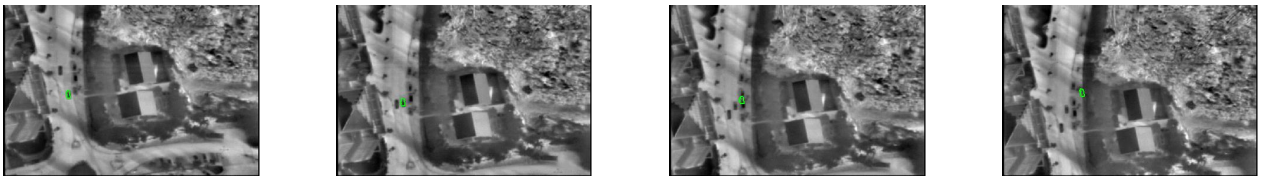

Figure 1. Loss of tracking in a clutter environment due to the multiple hypothesis nature of particle filtering caused by objects sharing similar statistical information.

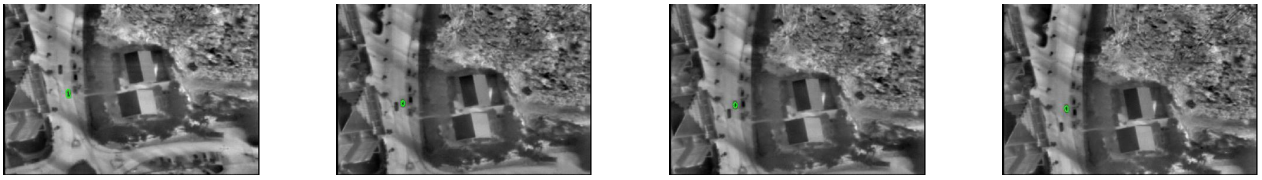

Figure 2. Successful result of the proposed algorithm in a clutter environment when a penalization scheme is introduced.

Figure 3. Several captures demonstrating the algorithm by tracking a person with a multimodal distribution in a crowded sequence.

Figure 4. Several captures demonstrating the proposed algorithm tracking under camera deformation as the object moves through the scene.

REFERENCES

1. A. Blake and M. Isard, *Active Contours*, Springer, Cambridge, 1998.
2. D. Terzopoulos and R. Szeliski, "Tracking with kalman snakes," *Active Vision* , pp. 3–20, 1992.
3. A. Yezzi and S. Soatto, "Deformation: Deforming motion, shape average and the joint registration and approximation of structures in images," *International Journal of Computer Vision* **53**.
4. T. Georgiou, "An intrinsic metric for power spectral density functions," *IEEE Trans. on Signal Processing Letters* (8), pp. 561–563, 2007.
5. T. Georgiou, "Distances and riemannian metrics for spectral density functions," *IEEE Trans. on Signal Processing* (8), pp. 3395–4003, 2007.
6. M. Isard and A. Blake, "Condensation - conditional density propagation for visual tracking," *International Journal of Computer Vision* **29**(1).
7. Y. Chen, T. Huang, and Y.Rui, "Parameteric contour tracking using unscented kalman filter," *Proceedings of the International Conference on Image Processing* **3**.
8. P. Li, T. Zhang, and B. Ma, "Unscented kalman filter for visual curve tracking," *Image and Vision Computing* **22**.
9. J. Sethian, *Level Set Methods and Fast Marching Methods, Second Edition*, Springer, New York, NY, 1999.
10. S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, Cambridge University Press, New York, NY, 2003.
11. T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Process.* **10**, pp. 266–277, Feb. 2001.
12. Y. Rathi, O. Michailovich, J. Malcolm, and A. Tannenbaum, "Seeing the unseen: Segmenting with distributions," in *Proc. Int. Conf. Sig. Imag. Proc.*, 2006.
13. S. Dambreville, M. Niethammer, A. Yezzi, and A. Tannenbaum, "A variational framework combining level-sets and thresholding," in *Proceedings in British Machine Vision Conference.*, 2007.
14. M. Rousson and R. Deriche, "A variational framework for active and adaptive segmentation of vector valued images," in *Wkshp. Motion Vid. Computi.*, p. 56, 2002.
15. N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, pp. 226–280, Mar. 2000.
16. M. Niethammer and A. Tannenbaum, "Dynamic geodesic snakes for visual tracking," *Proceedings in Computer Vision and Pattern Recognition* **1**, pp. 660–667, 2004.
17. Y. Rathi, N. Vaswani, A. Tannenbaum, and A. Yezzi, "Particle filtering for geometric active contours with application to tracking moving and deforming objects," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
18. J. Jackson, A. Yezzi, and S. Soatto, "Tracking deformable moving objects under severe occlusions," *Proceedings in Conference Decision and Control* , Dec. 2004.
19. D. Freedman and T. Zhang, "Active contours for tracking distributions," *IEEE Transactions on Image Processing* **13**(4), 2004.
20. Y. Rathi, N. Vaswani, A. Tannenbaum, and A. Yezzi, "Tracking deforming objects using particle filtering for geometric active contours," *IEEE Transactions on Pattern Analysis and Machine Interlligence (PAMI)* (8), 2007.
21. Y. Rathi, S. Dambreville, and A. Tannenbaum, "Particle filtering with dynamic shape priors," *International Conference on Image Analysis and Recognition* , 2006.
22. S. Dambreville, Y. Rathi, and A. Tannenbaum, "Tracking deforming objects with unscented kalman filtering and geometric active contours," *In American Control Conference on Image Analysis and Recognition* , 2006.
23. N. Gordon and D. Salmond and A. Smith, "Novel approach to nonlinear/nongaussian bayesian state estimation," *IEEE Proceedings-F (Radar and Signal Processing)* **140**(2), pp. 107–113, 1993.
24. B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, 2004.
25. A. Doucet, N. deFreitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*, Springer, 2001.
26. A. Doucet, S. Godsill, and C. Andrieu, "On sequential monte carlo sampling methods for bayesian filtering," *Statistics and Computing* **19**(3), pp. 197–208, 200.