

Model

Willie

September 2, 2011

The Data

Data are extracted from stationary, single-camera videos that contain moving objects. A moving object produces a trail of ‘motion points’, which are data representing groups of pixels that exhibit sufficient change between consecutive frames. The D -dimensional position of each motion point (where $D = 2$ for a typical video) is recorded, along with a V -dimensional vector representing a distribution over the motion point’s pixel color values. Hence, the data consists of motion point observations $x = \{x^{pos_1}, \dots, x^{pos_D}, x^{col_1}, \dots, x^{col_V}\}$.

Likelihood

For a given object k at video frame t , the set of positions of the object’s motion points are modeled with a product (over each position dimension) of one-dimensional Gaussians

$$P(x_{t_i}^{pos} | \mu_{t,k}, \lambda_{t,k}) = \prod_{d=1}^D \mathcal{N}(x_{t_i}^{pos_d} | \mu_{t,k}^d, \lambda_{t,k}^d) \quad (1)$$

where $x_{t_i}^{pos} = \{x_{t_i}^{pos_1}, \dots, x_{t_i}^{pos_D}\}$, $\mu_{t,k}^d$ is the mean of the k^{th} Gaussian at time t for dimension d , and $\lambda_{t,k}^d$ is the precision of the k^{th} Gaussian at time t for dimension d . Likewise, for a given object k at video frame t , the set of color vectors of the object’s motion points are modeled with a multinomial

$$P(x_{t_i}^{col} | m_{t,k}) = Mult(x_{t_i}^{col} | m_{t,k}) \quad (2)$$

where $x_{t_i}^{col} = \{x_{t_i}^{col_1}, \dots, x_{t_i}^{col_V}\}$, $m_{t,k} = \{m_{t,k}^1, \dots, m_{t,k}^V\}$, $\sum_{v=1}^V m_{t,k}^v = 1$, and $m_{t,k}^v > 0 \quad \forall v \in \{1, \dots, V\}$. The likelihood for a motion point observation is thus

$$P(x_{t_i} | \mu_{t,k}, \lambda_{t,k}, m_{t,k}) = Mult(x_{t_i}^{col} | m_{t,k}) \prod_{d=1}^D \mathcal{N}(x_{t_i}^{pos_d} | \mu_{t,k}^d, \lambda_{t,k}^d) \quad (3)$$

Base Distribution \mathbb{G}_0

\mathbb{G}_0 is the base distribution of the underlying time-dependent Dirichlet process mixture; it also serves as a prior distribution for the parameters present in the likelihood. We make use of conjugate priors in the base distribution to allow for more efficient computation. In

particular, for an object k at time t , we let the prior distribution over the parameters of the motion point position distribution, $\mu_{t,k}$ and $\lambda_{t,k}$, be

$$\mathbb{G}_0^{pos}(\mu_{t,k}, \lambda_{t,k} | \mu_0, n_0, a, b) = \prod_{d=1}^D [\mathcal{N}(\mu_{t,k}^d | \mu_0, n_0 \lambda_{t,k}^d) Ga(\lambda_{t,k}^d | a, b)] \quad (4)$$

where μ_0, n_0, a , and b are parameters of the base distribution. Furthermore, for object k at time t , we let the prior distribuion over the parameters of the motion point color vector distribution be

$$\mathbb{G}_0^{col}(m_{t,k} | q) = Dir(m_{t,k} | q) \quad (5)$$

where $q = \{q^1, \dots, q^V\}$ is a parameter of the base distribution, where $q^v > 0 \ \forall v \in \{1, \dots, V\}$. The base distribution is thus

$$\mathbb{G}_0(\mu_{t,k}, \lambda_{t,k}, m_{t,k} | \mu_0, n_0, a, b, q) = Dir(m_{t,k} | q) \prod_{d=1}^D [\mathcal{N}(\mu_{t,k}^d | \mu_0, n_0 \lambda_{t,k}^d) Ga(\lambda_{t,k}^d | a, b)] \quad (6)$$

Transition Kernels

Let $\phi_{t,k}^{pos} = \{\mu_{t,k}, \lambda_{t,k}\}$ and $\phi_{t,k}^{col} = \{m_{t,k}\}$. We define two transition kernels, $P(\phi_{t,k}^{pos} | \phi_{t-1,k}^{pos})$ and $P(\phi_{t,k}^{col} | \phi_{t-1,k}^{col})$, which dictate, respectively, the random walk of the motion point position parameters and the motion point color vector parameters over time. The transition kernels must be chosen so that their invariation distributions are \mathbb{G}_0 , i.e. such that the following hold:

$$\int \mathbb{G}_0^{pos}(\phi_{t-1,k}^{pos}) P(\phi_{t,k}^{pos} | \phi_{t-1,k}^{pos}) d\phi_{t-1,k}^{pos} = \mathbb{G}_0^{pos}(\phi_{t,k}^{pos}) \quad (7)$$

$$\int \mathbb{G}_0^{col}(\phi_{t-1,k}^{col}) P(\phi_{t,k}^{col} | \phi_{t-1,k}^{col}) d\phi_{t-1,k}^{col} = \mathbb{G}_0^{col}(\phi_{t,k}^{col}) \quad (8)$$