Note Title 11/12/2007

(1). Discrete Deformable Templates and Matching.

These deformable template models are defined on interest point features (sparse). Attributed Features (AF's)

$\{(x_i, A_i) : i = 1 \ldots N\}$

x_i – spatial position..
A_i – attributes.

EG. response of Gabors
Lowe's SIFT Features
Edgelets, ..

Basic assumption
→ some of the AF's correspond to features on the object
→ others correspond to background.

(2) **Object Model**

AF's $\{(y_a, B_a) : a = 1 \text{ to } M\}$

y_a - positions
B_a - attributes, or parameters defining attribute distribution.

Define a set of correspondence variables $\{V_{ai} : a = 1 \text{ to } M,\ i = 1 \text{ to } N\}$

$V_{ai} = 1$, if AF (y_a, B_a) on the model corresponds to AF (x_i, A_i) in the image.

$V_{ai} = 0$, otherwise.

Define a "junk node" $i = 0$
$V_{a0} = 1$, means that AF (y_a, B_a) on the model is unmatched.

Constraint: $\sum_{i=0}^{N} V_{ai} = 1, \ \forall a.$

T - unknown translation variable.

(3). <u>Define</u> a generative model for generating the image AF's.

$$P(\{(x_i, A_i)\} \mid \{V_{ai}\}, \{(y_a, B_a)\}, T)$$
$$= \frac{1}{Z_1} e^{-\sum_{a,i} V_{ai} \phi(x_i, y_a, A_i, B_a, T)}$$

<u>what</u> is $\phi(x_i, y_a, A_i, B_a, T)$?

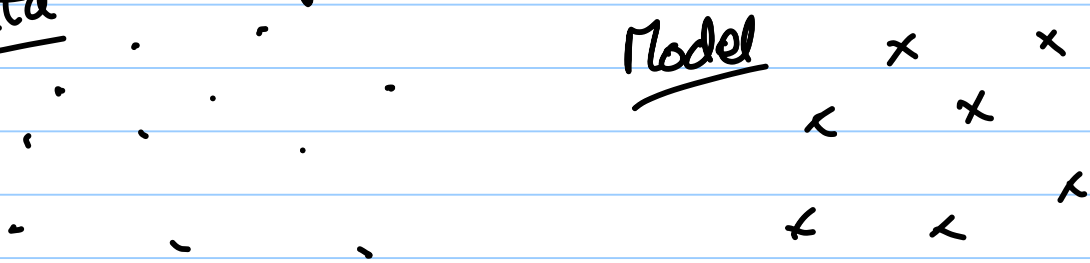
It depends on the formulation. Typically

$$\phi(x_i, y_a, A_i, B_a) = (x_i - y_a - T)^2 + (A_i - B_a)^2$$

<u>This</u> is a Gaussian distribution on spatial position and attributes.

→ e.g. B_a is the mean attribute value of the distribution.

<u>Data</u> <u>Model</u>



(4) Need a prior model for the $\{V_{ai}\}$.

Typically $P(\{V_{ai}\}) = \frac{1}{Z_2} e^{-\lambda \sum_a V_{a0}}$.

penalizes unmatched AF's on the object.

Can also have a prior model on T — but assume this to be a uniform distribution.

$$P(\{(x_i, A_i)\} \mid \{V_{ai}\}, \{(y_a, B_a)\}, T)\, P(\{V_{ai}\}).$$

Treat the correspondences $\{V_{ai}\}$ as hidden variables to be summed out.

$$P(\{(x_i, A_i)\} \mid \{(y_a, B_a)\}, T) = \sum_{\{V_{ai}\}} P(\{(x_i, A_i)\} \mid \{V_{ai}\}, \{(y_a, B_a)\}, T)\, P(\{V_{ai}\})$$

(5) To do inference, we have to estimate the translation T and deal with the hidden variables.

The E.M. algorithm is suitable for doing this

- but may get stuck in a local minima unless good initial conditions.

E.M. Algorithm.

Introduce distribution $q(V)$.

Free energy:

$$F[T, q(V)] = \sum_V q(V) \log q(V) - \sum_V q(V) \log P(X, A, V \mid T)$$

This can also be written as:

$$F[T, q(V)] = -\log P(X, A \mid T) + \sum_V q(V) \log \frac{q(V)}{P(V \mid X, A, T)}$$

(6) <u>Minimize</u> $F[T, q(V)]$ w.r.t. T and $q(V)$ in alteration.

$$q^{t+1}(V) = P(V|X, A, T^t).$$

$$T^{t+1} = \arg\min_{T} - \sum_{V} q^{t+1}(V) \log P(X, A, V|T)$$

<u>Correspondence</u> constraints.

$$\sum_{i=0}^{N} V_{ai} = 1, \quad \forall a.$$

If we impose these constraints, then the nature of the distribution $P(X, A|V, T)\, P(V)$ <u>implies</u> that $q^{t+1}(V)$ is a factorizable distribution

$$q(V) = \prod_{a} q_a(\underline{V}_a)$$

where $\underline{V}_a = (V_{a0}, V_{a1}, \ldots V_{aN})$

<u>correspondence</u> is independent over the object AF's

(7) What is the price for independence?

It means that we allow two AF's on the object to match to the same AF in the image

$\exists\, i$ s.t. $V_{ai} = 1$ & $V_{bi} = 1$

for some a & b.

This is unlikely to happen since it will have low probability (penalized by the ϕ term). Later, we will give another way to prevent it.

$$q_a(V_{ai}=1) = \frac{e^{-\phi(x_i, A_i, y_a, B_a, T)}}{\sum_{j=1}^{N} e^{-\phi(x_j, A_j, y_a, B_a, T)} + e^{-\lambda}}$$

$$q_a(V_{a0}=1) = \frac{e^{-\lambda}}{\sum_{j=1}^{N} e^{-\phi(x_j, A_j, y_a, B_a, T)} + e^{-\lambda}}$$

(8) $\hat{T} = \underset{T}{\mathrm{ARG\,MIN}} \left\{ -\sum_{V} q(V) \log P(X, A, V \mid T) \right\}$

$= \underset{T}{\mathrm{ARG\,MIN}} \sum_{i,a} q_a(V_{ai})\, \phi(X_i, A_i, y_a, B_a, T)$

If $\phi(X_i, A_i, y_a, B_a, T) = (X_i - y_a - T)^2 + (A_i - B_a)^2$

<u>Then</u> $\hat{T}$ can be solved by quadratic minimization:

When does this approach work?

The EM algorithm is fairly good if the number of object AF's is fairly similar to the number of image AF's.

<u>And</u> if the feature points have very different attributes. → (i.e. easy to distinguish one attribute from another).

(9) Otherwise, there are too many local minima and the correspondence problem (determining the $\{V_{ia}\}$) becomes too difficult.

There is a chicken & egg problem. If you can estimate the translation T, then it is often easy to estimate the correspondence $\{V\}$, And vice-versa.

I'll return to this issue in later lectures.

Variations on the model.

Allow the $\{y_a\}$ to become variables, remove the T, put a prior $P(\{y_a\})$ on the $\{y_a\}$'s.

E.G. $P(\{y_a\}) = \frac{1}{Z} e^{-E[\{y_a\}]}$

where $E[\{y_a\}]$ is a quadratic function of the differences of the $\{y_a\}$ $\to$ e.g. $\sum_{a,b}(y_a - y_b)^2$

(10) Repeat the same formulation

$$P(\{X,A\} \mid \{y,B\}, V)\ P(V)\ P(\{B\})$$

Free Energy

$$\sum_V q(V) \log q(V) - \sum_V q(V) \log P(\{B\}, V \mid \{X,A\})$$

$q(V)$ factorizable as before:

$$q_a(V_{ai}=1) = \frac{e^{-\phi(x_i, A_i, y_a, B_a)}}{\sum_{j=1}^{N} e^{-\phi(x_j, A_j, y_a, B_a)} + e^{-\lambda}}$$

$$q_a(V_{a0}=1) = \frac{e^{-\lambda}}{\sum_{j=1}^{N} e^{-\phi(x_j, A_j, y_a, B_a)} + e^{-\lambda}}$$

$$\hat{B} = \arg\min_{B} \sum_{i,a} q_a(V_{ai})\, \phi(x_i, A_i, y_a, B_a) + E_p[\{y_a\}]$$

Again, if $\phi(..)$ is quadratic in the $\{y_a\}$ and $E_p[\{y_a\}]$ is quadratic – then we can solve for $\hat{B}$ by quadratic minimization.

Beyond EM, Mean Field Theory.

Suppose we want to prevent two AF's in the object from matching the same AF in the image.

We can impose constraints by adding energy terms to the prior $P[V]$

$$E_V[V] = -\tau \sum_{a,b,i} V_{ai} V_{bi}.$$

(12) This will prevent the $q(.)$ from factorizing, because it includes terms that couple the matching of object AF a with object AF b.

<u>Instead</u> restrict the form of $q(\,)$ to ensure that it is factorizable. — this is an approximate to EM.

It can be shown that this corresponds to replacing the free energy by

$$\sum_{i,a} q_a(V_{ai}) \quad \phi(x_i, A_i, y_a, B_b) + E_p[\{y_a\}]$$

$$+ \sum_{i,a} q_a(V_{ai}) \log q(V_{ai})$$

with constraint. $\sum_{i=0}^{N} q_a(V_{ai}) = 1 \quad \forall a.$

Mean Field Approximation.

(13) Alternatively, we can restrict to a problem where $N=M$ and all points are matched.

Impose 1-1 constraints by

$$\sum_i V_{ai} = 1, \forall a \quad \& \quad \sum_a V_{ai} = 1, \forall i$$

Put into free energy with MFT approximation

$$\sum_{i,a} q_a(V_{ai})\, \phi(x_i, A_i, y_a, B_a) + E_p[\{y_a\}]$$

$$+ \sum_{i,a} q_a(V_{ai}) \log q(V_{ai})$$

$$+ \sum_i \lambda_i \{\sum_a q_a(V_{ai}) - 1\}$$

Lagrange multiplier.

$$+ \sum_a \mu_a (\sum_i q_a(V_{ai}) - 1)$$

Solving for the q → corresponds to the linear assignment problem – (polynomial).