Lecture 1) AAM's and FORMS. Note Title 3/1/2009

Active Appearance Models (AAM's)

– previous shape model (mask) is too simple. Doesn't allow flexibility.

Simplest Model:

$$I(\underline{x}) = T(\underline{\phi}(\underline{x}))$$

$\underline{\phi}(\underline{x})$ spatial warp.

where $\underline{\phi}(\underline{x}) = \underline{x} + \sum_{\mu} \alpha_{\mu} B_{\mu}(x)$

Basis Functions.

Can learn the Basis functions — and a prior $p(\alpha)$ on the coefficients — having a set of labelled datapoints (e.g. Cootes and Taylor).

(Alternatively, learn $\phi(\underline{x})$ from examples by assuming a form of smoothness – e.g. Hallinan).

The function $T(.)$ is the appearance model. It can be extended to

$$T(.) = \sum_i \beta_i \gamma_i(.)$$

– i.e. linear combination of basis functions

These linear combinations could occur by modeling the appearance of a class – e.g. faces.

Align faces (estimate $\phi(\underline{x})$ by hand). Then do PCA to estimate the $\{\gamma_i(.)\}$ and the $\{\beta_i\}$

Alternatively, for a single object – the $\gamma_i(.)$ can correspond to the basis functions for lighting.

(2)

Full model:

$$I(x) = \sum_i \beta_i \gamma_i \left(\underline{x} + \sum_\mu \alpha_\mu B_\mu(x) \right)$$

$\gamma_i(.)$ eigenvectors of appearance
$B_\mu(.)$ eigenvectors of spatial warps

Match the model to the image by least squares

$$E[\beta, \alpha; \gamma, B] = \sum_x \left\{ I(x) - \sum_i \beta_i \gamma_i \left(\underline{x} + \sum_\mu \alpha_\mu B_\mu(x) \right) \right\}^2$$

Generative Model:

$$P[I(x) \mid \beta, \alpha : \gamma, \beta] = \frac{1}{Z} e^{-E[\beta, \alpha; \gamma, B]}$$

Prior on $[\beta, \alpha]$

Note: this is one of the few generative models that really works – but it has limitations, can only deal with limited variations of appearance and shape deformations.

we return to the limitations for shape later.

Inference for an AAM.

Can be done by minimizing w.r.t. $\{\beta_i\}$ and $\{\alpha_\mu\}$ alternatively. Requires good initial conditions – or stuck in a local minima.

Min w.r.t. $\{\beta\}$ $\rightarrow$ solve linear equation (straightforward)

Min w.r.t. $\{\alpha\}$ $\rightarrow$ need to solve non-linear equation possible by steepest descent (at multiscale $\rightarrow$ blurring the images)

(Cootes & Taylor Handout)

(3) <u>How</u> to learn an AAM with limited supervision
→ see handout Kokkinos & Yuille.

Simplify the problem:
Perform Edge Detection and Ridge Detection
Blur edges — to give better domains of attraction for steepest descent.

$T(x)$ — blurred edge map.

Examples $\{I^\mu(x)\}$ — blurred edgemaps (cows, faces, hands)

<u>Perform</u> learning by the EM. algorithm

Parameters to be learnt. $T(.)$ $\{B_\mu(.)\}$
variances σ_μ^2 → priors on the (α's).
$\{\bar{\alpha}_\mu\}$

<u>Hidden</u> variables $\{\alpha_\mu\}$

$$P(\alpha) = \frac{1}{Z} e^{-\frac{(\alpha_\mu - \bar{\alpha}_\mu)^2}{2\sigma_\mu^2}}.$$

<u>Maximize</u> w.r.t. $T(.)$, $\{B_\mu(.)\}$.

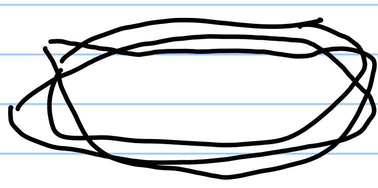
$$\prod_\mu P(I^\mu | \{B\}, T, \{\bar{\alpha}_\mu\}, \{\sigma_\mu\}) \, P(\{\alpha_\mu\} | \{\bar{\alpha}_\mu\})$$

<u>Difficult</u> to sum out the hidden variables. Instead maximize w.r.t. hidden variables

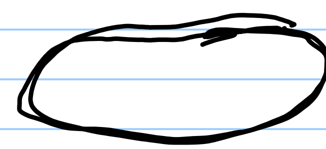
steepest descent at multiple scale.

Good initialization required.
Roughly align the objects.
Perform mean shift perpendicular to the contours.

Before mean shift.



After mean shift. (perp to contours)

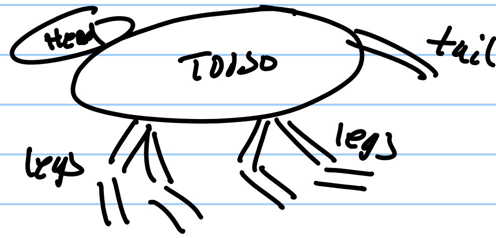


perpendicular mean shift — purpose is to take rough alignment and make it more precise.
Displacements caused by mean shift can initialize the PCA.

(4.) Limitations of AAM's

The torso can be well described by an AAM.

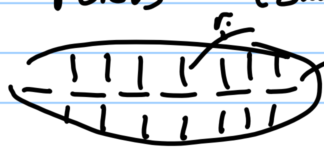
But the legs, tail, head have greater variability.



Represent object as a set of parts.

- Each part could be represented by an AAM (i.e. basis function determined by PCA).

E.G. FORMS (Zhu & Yuille).



vectors $\{r_i\}$

Do PCA on $\{r_i\}$

Objects represented by their expansions of $\{r_i\}$ in terms of basis function.

Plus spatial relationship between them.

Limitations of FORMS (and related systems).

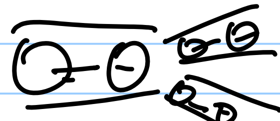
They use only silhouette information and no internal edges.

E.G. will not work well. →

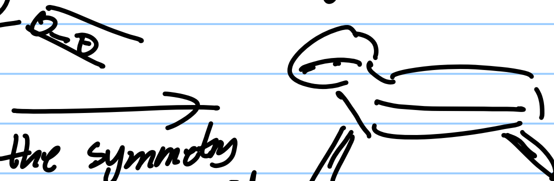
or

How to compute the representation?

Decompose objects by finding the symmetry axes

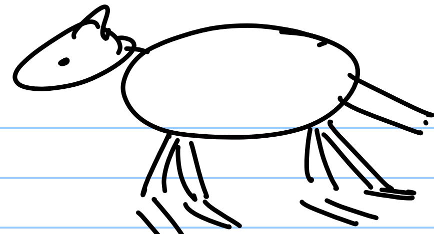


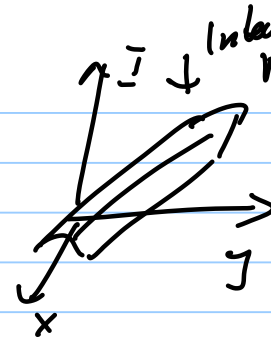
→ segment the object.



In practice, the symmetry axes only give an approximate decomposition of the object into parts — use as an initial estimate that can be improved by high level knowledge (FORMS)

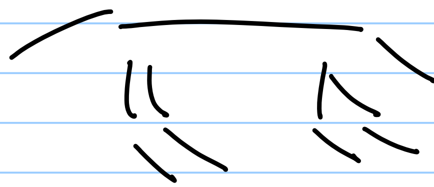
(5)





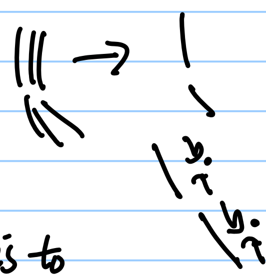
ridge detectors
can detect elongated shapes

Estimate number of parts from the ridge maps

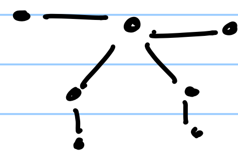


cluster on these using
mean-shift perp. to the line

Then do mean-shift in the
direction along the line



Then use this to
learn a model



each node represents a
part,
the edges specify the
spatial relations.