Correlation Based Matching

COMPSCI 773 S1 T

VISION GUIDED CONTROL

A/P Georgy Gimel'farb

Local and Global Optimisation

- 3-D reconstruction:
  - Search for the **max similarity** (min dissimilarity) between the corresponding regions or pixels in images of a stereo pair
  - **Basic problem**: an adequate measure for similarity (dissimilarity)
- Similarity (or dissimilarity) measure has to account for real image distortions due to separate image acquisition and different views:
  - Global and local uniform or non-uniform contrast / offset differences between corresponding signals
  - Geometric differences (projective distortions, partial occlusions)
- Similarity (or dissimilarity) has to include regularising constraints:
  - To deal with **partial occlusions** or **multiple equivalent** optima
  - For a *single continuous surface*: **visibility** and **ordering constraints**

COMPSCI 773 1

Note: scene differences due to acquisition at different time moments; occluded walls of high buildings; different brightness / contrast

COMPSCI 773 2

Local and Global Optimisation

- Possible scenarios of reconstructing a 3-D scene:
  1. Exhausting Δ^{NM} variants (Δ - range of disparities) of 3D surfaces in the space XYD of stereo disparities by global constrained optimisation

     $$\mathbf{d}^* = \max_{\mathbf{d}}\{\text{Similarity}(\mathbf{d}\,|\,\mathbf{g}_L,\mathbf{g}_R)\} \text{ or } \mathbf{d}^* = \min_{\mathbf{d}}\{\text{Dissimilarity}(\mathbf{d}\,|\,\mathbf{g}_L,\mathbf{g}_R)\}$$

     - Constraints on neighbouring disparities (to ensure smoothness and visibility of 3D surfaces)
  2. Independent selection of each 3-D point by local optimisation
     - May produce physically inconsistent surfaces (violating visibility constraints)
  3. Successive search by local optimisation for each next small surface patch in order to add it to the already found part of the goal surface
     - Typically, leads to ***accumulation of local errors*** and thus to arbitrarily large errors in a reconstructed surface

COMPSCI 773 3

Local Optimisation

- **Pros**:
  - Usually, quite simple computations
  - Easily accounts for both x- and y-disparities in the images
- **Cons** for independent selection of 3-D points:
  - Surfaces found may violate visibility and continuity constraints
- **Cons** for guiding each next search by the current surface:
  - Due to accumulation of local errors, search regions after a few steps may become completely wrong with no cues how to return to a true surface
- **Cons** in both cases:
  - Needs intensive on- or off-line editing of any reconstructed 3-D scene to fix inevitable large errors

COMPSCI 773 4

Simplifying Assumptions

- Elements to match by local correlation – relatively small image windows of fixed or (less frequently) adaptable size
- Pixel correspondence is given by the window position such that the similarity score within a search region is maximised
- Assumptions:
  - Canonical stereo geometry with parallel optical axes (no y-disparity)
  - Frontal (or slant) planar surfaces, at least, approximately
  - Simple models of signal deviations between the images:
    - Only uniform independent Gaussian or any central-symmetric noise
    - Uniform contrast and offset distortions in addition to such a noise

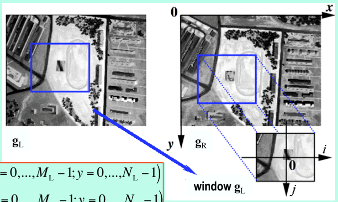
COMPSCI 773 5

Correspondence by Correlation

Matching rectangular windows of size $(2a+1)\times(2b+1)$ representing a frontal planar surface

Window in $\mathbf{g}_L$: $\left(g_{L:x'+i,y'+j} : i=-a,\dots,0,\dots,a; \; j=-b,\dots,0,\dots,b\right)$

Window in $\mathbf{g}_R$: $\left(g_{R:x+i,y+j} : i=-a,\dots,0,\dots,a; \; j=-b,\dots,0,\dots,b\right)$



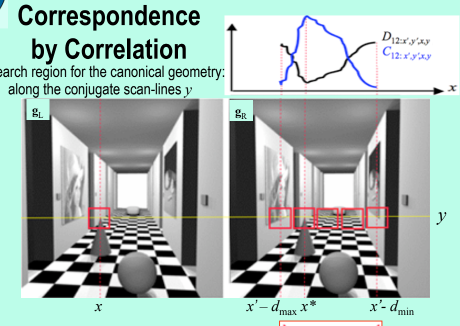
$\mathbf{g}_L = \left(g_{L:x,y} : x=0,\dots,M_L-1; y=0,\dots,N_L-1\right)$

$\mathbf{g}_R = \left(g_{R:x,y} : x=0,\dots,M_R-1; y=0,\dots,N_R-1\right)$

COMPSCI 773 6

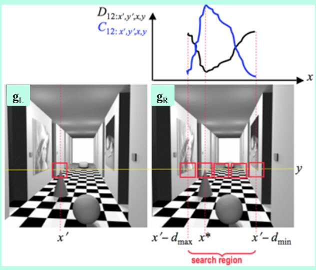
Correspondence by Correlation

Search region for the canonical geometry: along the conjugate scan-lines y



COMPSCI 773 7

Correspondence by Correlation



For any point x' in $\mathbf{g}_L$, the corresponding point (x^*,y) in $\mathbf{g}_R$ minimises the Euclidean distance:

$$x^* = \arg\min_{x\in SR} D_{LR:x',y;x,y}$$

or maximises the correlation:

$$x^* = \arg\max_{x\in SR} C_{LR:x',y;x,y}$$

between the fixed window in $\mathbf{g}_L$ and the candidate window being moved across $\mathbf{g}_R$

COMPSCI 773 8

Similarity (Dissimilarity) Score

- The **similarity** or **dissimilarity score**: by minimising the distance between the corresponding signals in the windows with respect to allowable relative distortion
- Minimum squared distance ⇔ the maximal probability of signal matching for the uniform Gaussian noise
  - Actually, any central-symmetric noise such that its probability distribution monotonously decreases with the squared noise
- Minimum absolute distance if the probability distribution monotonously decreases with the absolute noise

COMPSCI 773 9

Symmetric Canonical Geometry

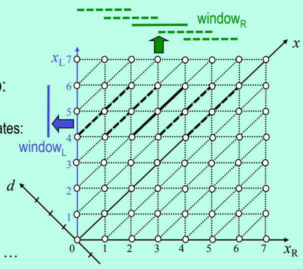
- Symmetric coordinates: $[x,y] \leftrightarrow [x_L,y], [x_R,y]$
  - Disparity: $d = x_L - x_R$
- Cyclopean disparity map: (x,y,d)
  - Symmetric (x,d)-coordinates:

$$\begin{pmatrix} x = (x_L + x_R)/2 \\ d = x_L - x_R \end{pmatrix} \leftrightarrow \begin{pmatrix} x_L = x + d/2 \\ x_R = x - d/2 \end{pmatrix}$$

$x_L, x_R = 0, 1, 2, \dots$

$x = 0, \frac{1}{2}, 1, 1\frac{1}{2}, 2, \dots$

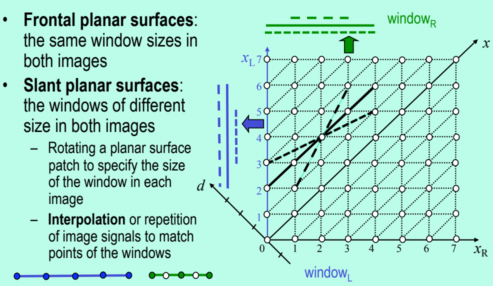
$d = \dots, -2, -1, 0, 1, 2, \dots$



COMPSCI 773 10

Matching Slant Planar Surfaces

- **Frontal planar surfaces**: the same window sizes in both images
- **Slant planar surfaces**: the windows of different size in both images
  - Rotating a planar surface patch to specify the size of the window in each image
  - **Interpolation** or repetition of image signals to match points of the windows



COMPSCI 773 11

Simple Signal Models: 1

- No local geometric distortion (in windows of fixed size)
  - i.e. patches of a frontal planar 3-D surface
- No photometric distortion except from a random noise:
  - Independent centred pixel-wise noise n with a fixed variance

$$g_{\mathrm{L}:x+\frac{d}{2},y} = g_{x,y} + n_{\mathrm{L}:x+\frac{d}{2},y};\quad g_{\mathrm{R}:x-\frac{d}{2},y} = g_{x,y} + n_{\mathrm{R}:x-\frac{d}{2},y}$$

$$\Rightarrow g_{\mathrm{L}:x+\frac{d}{2},y} = g_{\mathrm{R}:x-\frac{d}{2},y} + n_{x,y,d}$$

$$\Rightarrow D_{x,y,d} = \begin{cases} \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left|g_{\mathrm{L}:x+\frac{d}{2}+i,y+j} - g_{\mathrm{R}:x-\frac{d}{2}+i,y+j}\right| \\ \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x+\frac{d}{2}+i,y+j} - g_{\mathrm{R}:x-\frac{d}{2}+i,y+j}\right)^2 \end{cases}$$

Unknown signal for the 3-D point

Independent random pixel-wise noise with the same distribution

Independent random pixel-wise noise with the same distribution: $n_{x,y,d} = n_{\mathrm{L}:x+\frac{d}{2},y} - n_{\mathrm{R}:x-\frac{d}{2},y}$

COMPSCI 773 12

SSD / SAD Based Matching

- **Sum of squared distances** (SSD):

Symmetric cyclopean case: $D_{x,y,d} = \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x+\frac{d}{2}+i,y+j} - g_{\mathrm{R}:x-\frac{d}{2}+i,y+j}\right)^2$

Asymmetric conventional case: $D_{x',y';x,y} = \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x'+i,y'+j} - g_{\mathrm{R}:x+i,y+j}\right)^2$

- **Sum of absolute distances** (SAD):

Symmetric cyclopean case: $D_{x,y,d} = \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left|g_{\mathrm{L}:x+\frac{d}{2}+i,y+j} - g_{\mathrm{R}:x-\frac{d}{2}+i,y+j}\right|$

Asymmetric conventional case: $D_{x',y';x,y} = \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left|g_{\mathrm{L}:x'+i,y'+j} - g_{\mathrm{R}:x+i,y+j}\right|$

COMPSCI 773 13

Simple Signal Models: 2

- No local geometric distortions (in windows of fixed size)
  - i.e. a patch of a frontal planal 3-D surface
- Uniform contrast and offset photometric distortions
  - Independent centred noise n with a fixed variance
  - Fixed contrast factor α and offset β for every window
- Asymmetric left-to-right signal model:

$$g_{\mathrm{L}:x',y'} = \alpha g_{\mathrm{R}:x,y} + \beta + n_{x,y} \Rightarrow D_{x',y';x,y} = \min_{\alpha,\beta} D_{x',y';x,y;\alpha,\beta}$$

$$= \min_{\alpha,\beta}\sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x'+i,y'+j} - \alpha g_{\mathrm{R}:x+i,y+j} - \beta\right)^2$$

COMPSCI 773 14

Deriving the Correlation Score

$$\min_{\alpha,\beta} D_{x',y';x,y;\alpha,\beta} = \min_{\alpha,\beta}\sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x'+i,y'+j} - \alpha g_{\mathrm{R}:x+i,y+j} - \beta\right)^2 \Rightarrow \begin{cases} \left.\frac{\partial D_{x',y';x,y;\alpha,\beta}}{\partial\beta}\right|_{\alpha=\alpha^*,\beta=\beta^*} = 0 \\ \left.\frac{\partial D_{x',y';x,y;\alpha,\beta}}{\partial a}\right|_{\alpha=\alpha^*,\beta=\beta^*} = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x'+i,y'+j} - \alpha^* g_{\mathrm{R}:x+i,y+j} - \beta^*\right) = 0 \\ \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x'+i,y'+j} - \alpha^* g_{\mathrm{R}:x+i,y+j} - \beta^*\right) g_{\mathrm{R}:x+i,y+j} = 0 \end{cases} \Rightarrow \begin{cases} \beta^* = \frac{\sum_{i=-a}^{a}\sum_{j=-b}^{b} g_{\mathrm{L}:x'+i,y'+j} - \alpha^*\sum_{i=-a}^{a}\sum_{j=-b}^{b} g_{\mathrm{R}:x+i,y+j}}{(2a+1)(2b+1)} \\ \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x'+i,y'+j} - \alpha^* g_{\mathrm{R}:x+i,y+j} - \beta^*\right) g_{\mathrm{R}:x+i,y+j} = 0 \end{cases}$$

COMPSCI 773 15

Deriving the Correlation Score

Therefore, $\beta^* = m_{\mathrm{L}:x',y'} - \alpha^* m_{\mathrm{R}:x,y};\quad \alpha^* = \frac{S_{\mathrm{LR}:x',y';x,y}}{S_{\mathrm{RR}:x',y';x,y}}$ where

$m_{\mathrm{L}:x',y'} = \frac{S_{\mathrm{L}:x',y'}}{S_0};\ m_{\mathrm{R}:x,y} = \frac{S_{\mathrm{R}:x,y}}{S_0};$ Mean signals in the windows

$S_{\mathrm{L}:x',y'} = \sum_{i=-a}^{a}\sum_{j=-b}^{b} g_{\mathrm{L}:x'+i,y'+j};\ S_{\mathrm{R}:x,y} = \sum_{i=-a}^{a}\sum_{j=-b}^{b} g_{\mathrm{R}:x+i,y+j};\ S_0 = (2a+1)(2b+1)$ Window size

$S_{\mathrm{LR}:x',y';x,y} = \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x'+i,y'+j} - m_{\mathrm{L}:x',y'}\right)\left(g_{\mathrm{R}:x+i,y+j} - m_{\mathrm{R}:x,y}\right)$ Cross-product of the centred signals in the windows

$S_{\mathrm{RR}:x,y} = \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{R}:x+i,y+j} - m_{\mathrm{R}:x,y}\right)^2$ Non-normalised signal variance in the window on $\mathbf{g}_{\mathrm{R}}$

COMPSCI 773 16

Deriving the Correlation Score

- Minimum distance between the signals in the windows under the optimum relative contrast α^* and offset β^*:

$$D_{x',y';x,y} = S_{\mathrm{LL}:x',y'} - \frac{S^2_{\mathrm{LR}:x',y';x,y}}{S_{\mathrm{RR}:x,y}} = S_{\mathrm{LL}:x',y'}\left(1 - \frac{S^2_{\mathrm{LR}:x',y';x,y}}{S_{\mathrm{LL}:x',y'}S_{\mathrm{RR}:x,y}}\right)$$

where $S_{\mathrm{LL}:x',y'} = \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{\mathrm{L}:x'+i,y'+j} - m_{\mathrm{L}:x',y'}\right)^2$ Non-normalised signal variance in the window on $\mathbf{g}_{\mathrm{L}}$

$$D_{x',y';x,y} = S_{\mathrm{LL}:x',y'}\left(1 - C^2_{x',y';x,y}\right) \in [0, S_{\mathrm{LL}:x',y'}]$$

$$C_{x',y';x,y} = S_{\mathrm{LR}:x',y';x,y}\Big/\sqrt{S_{\mathrm{LL}:x',y'}S_{\mathrm{RR}:x,y}} \in [-1, 1]$$

Cross-correlation

COMPSCI 773 17

Correlation Matching

- Cross-correlation $C_{x',y';x,y}$ is most frequently used as the similarity measure (rather than the distance $D_{x',y';x,y}$ as the dissimilarity measure)
- Individual correlation search for a point corresponding to (x',y') in $\mathbf{g}_L$ in a search region SR:

$$(x^*,y^*) = \arg\min_{(x,y)\in \mathrm{SR}} D_{x',y';x,y} \quad \text{or} \quad (x^*,y^*) = \arg\max_{(x,y)\in \mathrm{SR}} C_{x',y';x,y}$$

  - The disparity vector for (x',y') in $\mathbf{g}_L$:

$$\mathbf{d}_{x',y'} = [d_{x',y'} = x' - x^*,\ \delta_{x',y'} = y' - y^*]$$

COMPSCI 773 18

Constraints on Contrast

- Relative contrast α^* and signal variances $S_{LL:x',y'}$ and $S_{RR:x,y}$ should be restricted
  - e.g. $\alpha_{min} \le \alpha^* \le \alpha_{max}$ and $S_{LL:x',y'} \ge \theta > 0$; $S_{RR:x,y} \ge \theta > 0$ to exclude inadequate matches of almost uniform windows

$$D_{x',y',x,y} = \begin{cases} S_{LL:x',y'} - \dfrac{S^2_{LR:x',y',x,y}}{S_{RR:x,y}} + S_{RR:x,y}\left(\alpha_{min} - \dfrac{S_{LR:x',y',x,y}}{S_{RR:x,y}}\right)^2 & \text{if} \quad \dfrac{S_{12:x',y',x,y}}{S_{22:x,y}} < \alpha_{min} \\ S_{LL:x',y'} - \dfrac{S^2_{LR:x',y',x,y}}{S_{RR:x,y}} & \text{if} \quad \alpha_{min} \le \dfrac{S_{12:x',y',x,y}}{S_{22:x,y}} \le \alpha_{max} \\ S_{LL:x',y'} - \dfrac{S^2_{LR:x',y',x,y}}{S_{RR:x,y}} + S_{RR:x,y}\left(\dfrac{S_{LR:x',y',x,y}}{S_{RR:x,y}} - \alpha_{max}\right)^2 & \text{if} \quad \alpha_{max} < \dfrac{S_{12:x',y',x,y}}{S_{22:x,y}} \end{cases}$$

COMPSCI 773 19

Fast Implementation

- Window-size independent sums $Q_{x,y} = \sum_{i=-a}^{a}\sum_{j=-b}^{b} q_{x+i,y+j}$
  - Accumulator: $\mathbf{Q}^{acc} = \left\{ Q^{acc}_{\xi,\eta} = \sum_{x=0}^{\xi}\sum_{y=0}^{\eta} q_{x,y} \,\middle|\, x = 0,\dots,M-1; y = 0,\dots,N-1 \right\}$

$$\Leftrightarrow Q^{acc}_{\xi,\eta} = q_{\xi,\eta} + Q^{acc}_{\xi-1,\eta} + Q^{acc}_{\xi,\eta-1} - Q^{acc}_{\xi-1,\eta-1}$$

  - Window sum: $Q_{x,y} = Q^{acc}_{x+a,y+b} - Q^{acc}_{x-a-1,y+b} - Q^{acc}_{x+a,y-b-1} + Q^{acc}_{x-a-1,y-b-1}$
- Straightforward correlation matching: $O(MNab\Delta)$
- Fast correlation matching: $O(MN\Delta)$
  - Δ - disparity range; ab - the dominant term of the window size
    - Even for a small window 11 x 11 pixels: ~ 120 times faster!

COMPSCI 773 20

Window Sums

Using the accumulated sums: $62-40-23+15 = \mathbf{14}$
$0+0+1+2+0+0+1+1+0+0+1+2+0+0+2+1+0+0+1+2 = \mathbf{14}$

2	2	1	0	0	0	0	1	1	2
2	2	1	0	0	0	0	1	2	1
2	2	1	0	0	0	0	2	1	1
2	2	1	0	0	0	0	1	2	1
2	2	1	0	0	0	0	1	1	2
2	2	1	0	0	0	0	1	2	1
2	2	1	0	0	0	0	2	1	1
2	2	1	0	0	0	0	1	2	1
2	2	1	0	0	0	0	1	1	2
2	2	1	0	0	0	0	1	2	1

2	4	5	5	5	5	5	6	7	9
4	8	10	10	10	10	10	12	15	18
6	12	15	15	15	15	15	19	23	27
8	16	20	20	20	20	20	25	31	36
10	20	25	25	25	25	25	31	38	45
12	24	30	30	30	30	30	37	46	54
14	28	35	35	35	35	35	44	54	63
16	32	40	40	40	40	40	50	62	72
18	36	45	45	45	45	45	56	69	80
20	40	50	50	50	50	50	62	77	89

COMPSCI 773 21

Fast Implementation

- $S_{L:x',y'}$ and $S_{R:x,y}$: $q_{x,y} = g_{L:x,y}$ or $g_{R:x,y}$
- $S_{LL:x',y'}$ and $S_{RR:x,y}$: $q_{x,y} = g^2_{L:x,y}$ or $g^2_{R:x,y}$

$$S_{LL:x',y'} = \sum_{i=-a}^{a}\sum_{j=-b}^{b}\left(g_{L:x'+i,y'+j} - m_{L:x',y'}\right)^2 = \sum_{i=-a}^{a}\sum_{j=-b}^{b} g^2_{L:x'+i,y'+j} - S_0 m^2_{L:x',y'}$$

- $S_{LR:x',y';x,y}$: using an accumulator for each current pair of disparities $d = x'-x;\ \delta = y'-y$
  - Compute the correlation (or distance) for every window under fixed disparities $(d,\ \delta)$ and keep in each location (x',y') both its current optimum and the disparity pair for it

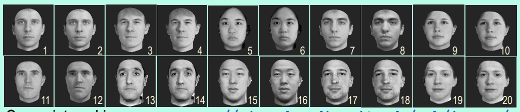
COMPSCI 773 22

Soft Least-squares Masking

- Least-squares image matching frequently fails in the presence of sizeable areas with large differences between corresponding signals
  - The differing signals are frequently called **outliers**
- We attempt to eliminate outliers by **soft masking** of suspicious pixels in an input image
  - Image differences are modelled with a mixture of a "valid" random noise and outliers
  - **Soft masks** with weights in [0.0,1.0] for each pixel: by an Expectation-Maximisation (EM) procedure

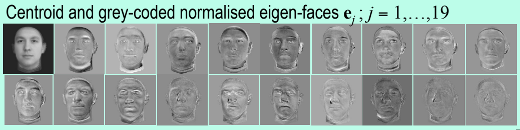
COMPSCI 773 23

MIT Face Database Subset



Co-registered images //vismod.media.mit.edu/pub/images/

Centroid and grey-coded normalised eigen-faces $\mathbf{e}_j$; $j = 1,\ldots,19$



COMPSCI 773 24

MIT Face Database Subset

- Ten subjects (20 co-registered 200 x 200 images)
- Two images of every subject under different illumination
  - Eigen-values λ_j and cumulative relative variances of signals $\nu_j = \sum_{i=1}^{j} \lambda_i \Big/ \sum_{i=1}^{19} \lambda_i$ represented by the j top eigen-faces:

	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_4$	$\mathbf{e}_5$	$\mathbf{e}_6$	$\mathbf{e}_7$	$\mathbf{e}_8$	$\mathbf{e}_9$
$\lambda_j \times 10^{-7}$	4.5	1.4	0.79	0.64	0.40	0.33	0.27	0.22	0.15
ν_j(%)	50	67	76	83	87	91	94	96	98

COMPSCI 773 25

Error probability mixture model

$$\Pr(\varepsilon_i)=\rho N(\varepsilon_i|\sigma) + (1-\rho)U(\varepsilon_i);\ \Pr(\boldsymbol{\varepsilon})=\Pr(\varepsilon_1)\cdot\ldots\cdot\Pr(\varepsilon_p)$$

implies that noise can be separated from outliers by masking out pixels with significantly larger than expected errors

Additional notation:

Q - the number of grey levels (typically, $Q = 256$)

Discrete errors $\varepsilon \in E = \{-Q+1, \ldots, -1, 0, 1, \ldots, Q-1\}$

ρ - unknown prior probability of non-outliers

$U(\varepsilon)$ - uniform, by assumption, distribution of outlying errors: $U(\varepsilon) = \frac{1}{2Q-1}$

$N(\varepsilon|\sigma)$ - discrete truncated zero-centred Gaussian distribution:

$$N(\varepsilon|\sigma) = \frac{1}{Z_\sigma}\exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right) \quad \text{where} \quad Z_\sigma = \sum_{\delta=-Q+1}^{Q-1}\exp\left(-\frac{\delta^2}{2\sigma^2}\right)$$

COMPSCI 773 26

Robust Symmetric Matching

- $\tilde{\mathbf{g}}_1, \tilde{\mathbf{g}}_2$: images derived from an unknown template $\tilde{\mathbf{g}}$ by varying uniform contrast (a_1 and a_2) and offset (b_1 and b_2) values; $\theta = (a_1, a_2, b_1, b_2, \tilde{\mathbf{g}})$
- Images are perturbed by independent per pixel errors ε caused by **noise** or **outlier**: for noise only, $\tilde{g}_{1i} = a_1\tilde{g}_i + b_1 + \varepsilon_{1i}$; $\tilde{g}_{2i} = a_2\tilde{g}_i + b_2 + \varepsilon_{2i}$
- **Maximum likelihood signal dissimilarity** of images in the presence of noise and outliers:

$$D_{12} = \min_\theta\left(-\ln\Pr(\tilde{\mathbf{g}}_1,\tilde{\mathbf{g}}_2)\right) = -\sum_{i=1}^{p}\left[\gamma_i \ln N(\varepsilon_i|\sigma) + (1-\gamma_i)\ln U(\varepsilon_i)\right]$$

$$= \frac{1}{2\sigma^2}\boldsymbol{\Phi}_{12} + \nu\ln(Z_\sigma) + (p-\nu)\ln(2Q-1) \quad \text{where} \quad \nu = \sum_{i=1}^{p}\gamma_i$$

COMPSCI 773 27

Robust Symmetric Matching

- Φ_{12} - minimum total squared error with respect to model parameters θ:

$$\boldsymbol{\Phi}_{12} = \min_\theta \sum_{i=1}^{p}\gamma_i\left(\varepsilon_{1i}^2 + \varepsilon_{2i}^2\right) = \tfrac{1}{2}\left[S_{11} + S_{22} - \sqrt{(S_{11}-S_{22})^2 + 4S_{12}^2}\right]$$

where

$$S_{kl} = \sum_{i=1}^{p}\gamma_i\left(\tilde{g}_{ki}-\mu_k\right)\left(\tilde{g}_{li}-\mu_l\right);\ k,l = 1,2;\ \mu_k = \tfrac{1}{\nu}\sum_{i=1}^{p}\gamma_i\tilde{g}_{ki}$$

- $\varepsilon_i = \alpha_2(g_{1i} - \mu_1) - \alpha_1(g_{2i} - \mu_2)$ - residual per pixel matching error after the ML estimation and elimination of model parameters:

$$\alpha_1^2 = \frac{1}{2}\left(1 + \frac{S_{11}-S_{22}}{\sqrt{(S_{11}-S_{22})^2 + 4S_{12}^2}}\right);\quad \alpha_2^2 = \frac{1}{2}\left(1 - \frac{S_{11}-S_{22}}{\sqrt{(S_{11}-S_{22})^2 + 4S_{12}^2}}\right)$$

COMPSCI 773 28

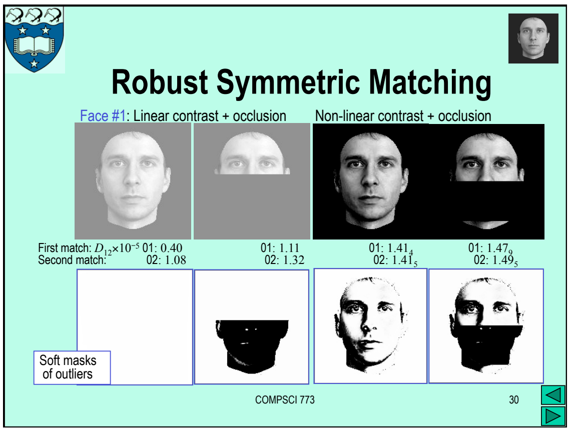
Robust Symmetric Matching

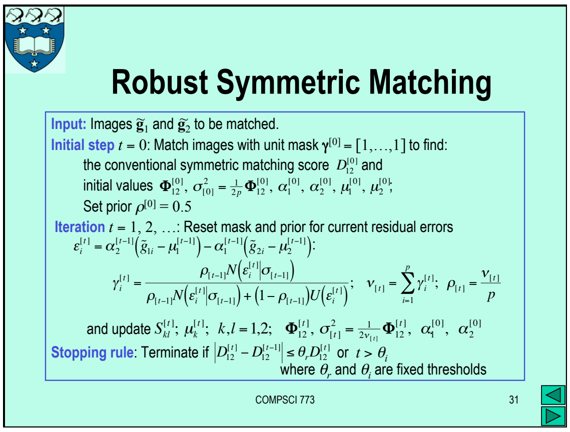
- Estimated noise variance:

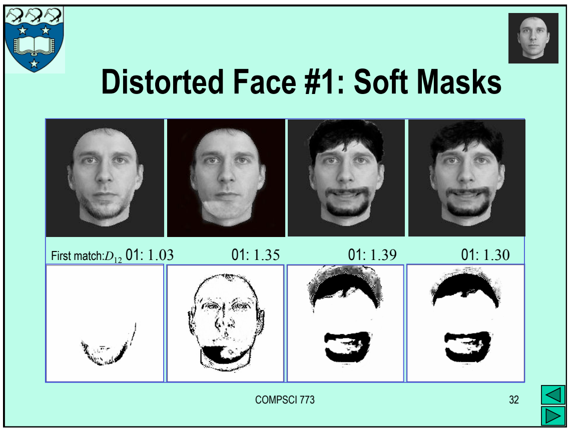
$$\sigma^2 = \frac{\boldsymbol{\Phi}_{12}}{2\nu} \quad \text{so that} \quad D_{12} = \nu\ln\left(1 + \sum_{\delta=1}^{Q-1}2\exp\left(-\frac{\delta^2\nu}{\boldsymbol{\Phi}_{12}}\right)\right) + \nu + (p-\nu)\ln(2Q-1)$$

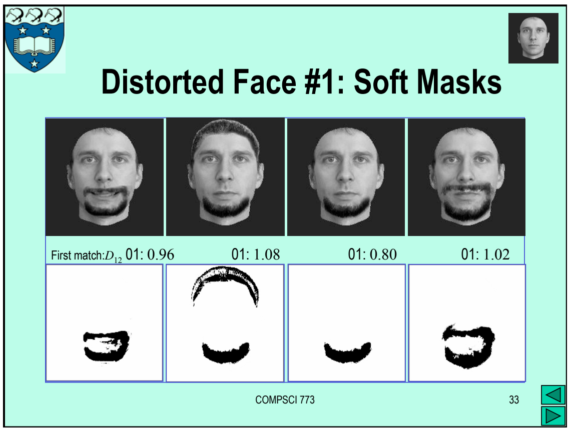
- Local minimum of Φ_{12} :
  - by EM-based iterative procedure
    - Re-evaluates soft masks $\boldsymbol{\gamma}$ and model parameters θ, similar to the robust PCA-based reconstruction algorithm

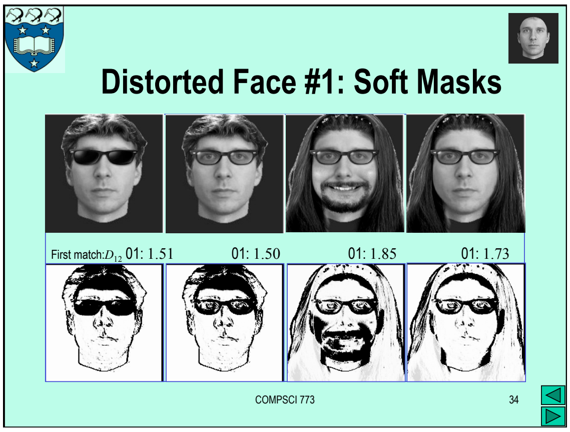
COMPSCI 773 29

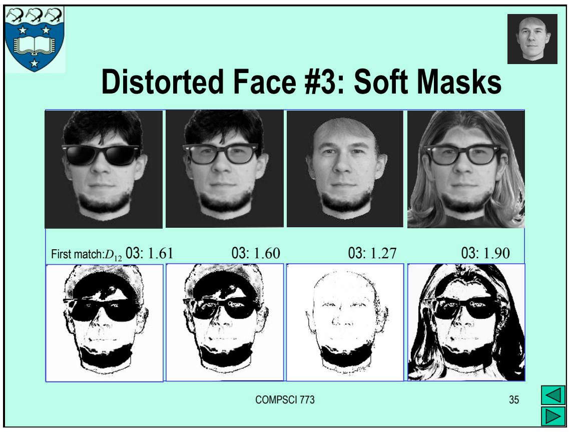

Robust Symmetric Matching
Face #1: Linear contrast + occlusion
Non-linear contrast + occlusion
First match: $D_{12}\times10^{-5}$ 01: 0.40
Second match: 02: 1.08
01: 1.11
02: 1.32
01: 1.41_4
02: 1.41_5
01: 1.47_9
02: 1.49_5
Soft masks of outliers
COMPSCI 773
30


Robust Symmetric Matching
Input: Images $\tilde{\mathbf{g}}_1$ and $\tilde{\mathbf{g}}_2$ to be matched.
Initial step $t = 0$: Match images with unit mask $\boldsymbol{\gamma}^{[0]} = [1,\ldots,1]$ to find:
the conventional symmetric matching score $D_{12}^{[0]}$ and
initial values $\boldsymbol{\Phi}_{12}^{[0]},\ \sigma_{[0]}^2 = \frac{1}{2p}\boldsymbol{\Phi}_{12}^{[0]},\ \alpha_1^{[0]},\ \alpha_2^{[0]},\ \mu_1^{[0]},\ \mu_2^{[0]}$;
Set prior $\rho^{[0]} = 0.5$
Iteration $t = 1, 2, \ldots$: Reset mask and prior for current residual errors
$\varepsilon_i^{[t]} = \alpha_2^{[t-1]}\left(\tilde{g}_{1i} - \mu_1^{[t-1]}\right) - \alpha_1^{[t-1]}\left(\tilde{g}_{2i} - \mu_2^{[t-1]}\right)$:
$\gamma_i^{[t]} = \frac{\rho_{[t-1]}N\left(\varepsilon_i^{[t]}\middle|\sigma_{[t-1]}\right)}{\rho_{[t-1]}N\left(\varepsilon_i^{[t]}\middle|\sigma_{[t-1]}\right) + \left(1-\rho_{[t-1]}\right)U\left(\varepsilon_i^{[t]}\right)};\quad \nu_{[t]} = \sum_{i=1}^{p}\gamma_i^{[t]};\quad \rho_{[t]} = \frac{\nu_{[t]}}{p}$
and update $S_{kl}^{[t]};\ \mu_k^{[t]};\ k,l = 1,2;\quad \boldsymbol{\Phi}_{12}^{[t]},\ \sigma_{[t]}^2 = \frac{1}{2\nu_{[t]}}\boldsymbol{\Phi}_{12}^{[t]},\ \alpha_1^{[0]},\ \alpha_2^{[0]}$
Stopping rule: Terminate if $\left|D_{12}^{[t]} - D_{12}^{[t-1]}\right| \le \theta_r D_{12}^{[t]}$ or $t > \theta_i$
where θ_r and θ_i are fixed thresholds
COMPSCI 773
31

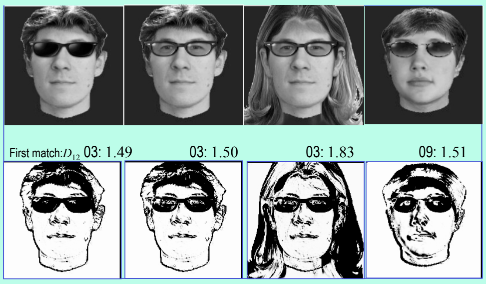

Distorted Face #1: Soft Masks
First match: D_{12} 01: 1.03
01: 1.35
01: 1.39
01: 1.30
COMPSCI 773
32


Distorted Face #1: Soft Masks
First match: D_{12} 01: 0.96
01: 1.08
01: 0.80
01: 1.02
COMPSCI 773
33


Distorted Face #1: Soft Masks
First match: D_{12} 01: 1.51
01: 1.50
01: 1.85
01: 1.73
COMPSCI 773
34


Distorted Face #3: Soft Masks
First match: D_{12} 03: 1.61
03: 1.60
03: 1.27
03: 1.90
COMPSCI 773
35

Distorted Faces #3,9: Soft Masks



COMPSCI 773 36

Distorted Face #9: Soft Masks

COMPSCI 773 37

Soft Masking: Matching Results

- Soft masking of outliers significantly increases the accuracy of image matching

Error rates for the symmetric LS matching:	*Conventional*	*Robust*
24 distorted MIT FDB faces (Slides 19 - 24)	54%	0%
40 half-occluded MIT FDB faces (*black top or bottom half*)	47%	5%

  - If models account for contrast and offset deviations and outliers, images can be matched or reconstructed more effectively even when large image regions are modified
- ***Directions of subsequent research***:
  - More adequate probability models of noise and outliers
  - Embedding the EM-based matching algorithms into a general least squares framework to account for geometric deviations

COMPSCI 773 38

Least-Squares Correlation

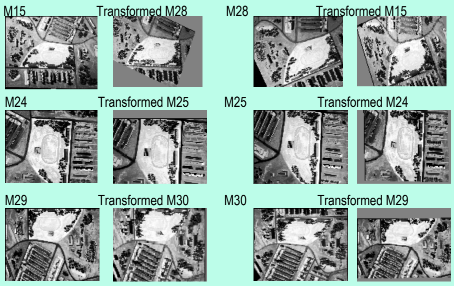
- Search for geometric transformations $\mathbf{t}^*$ maximising the cross-correlation matching score between the windows:

$$C_{\mathbf{t}^*} = \max_{\mathbf{t}}\{C_{\mathbf{t}}\}$$

- Simplified case: affine transformations $\begin{cases} x_R = a_1x_L + a_2y_L + a_3 \\ y_R = a_4x_L + a_5y_L + a_6 \end{cases}$
  - Combined exhaustive and directed (e.g. gradient based) search for affine parameters
  - Exhaustion of a sparse grid of the relative translations a_3 and a_5 of the fixed window g_1 with respect to the other image g_2
  - Directed unconstrained optimisation of $C_{\mathbf{t}}$ by all $\mathbf{t} = [a_1,\ldots,a_6]$ affine parameters starting from each grid point $\mathbf{t} = [1,0,a_3,0,1,a_6]$

COMPSCI 773 39

Affinely Transformed Windows



COMPSCI 773 40