Incremental Subclass Discriminant Analysis: A Case Study in Face Recognition

Hemank Lamba, Tejas Indulal Dhamecha, Mayank Vatsa, Richa Singh

Indraprastha Institute of Information Technology, Delhi, India

Monday, October 1, 2012
1 Motivation
   - Discriminant Analysis Techniques
   - Requirement of Incremental Techniques

2 Literature
   - Incremental PCA
   - Extension of IPCA to ILDA

3 Proposed Approach
   - Progression
   - Incremental SDA

4 Experiments and Analysis
   - Protocol
   - Performance

5 Conclusion and Future work
1 Motivation
   - Discriminant Analysis Techniques
   - Requirement of Incremental Techniques

2 Literature
   - Incremental PCA
   - Extension of IPCA to ILDA

3 Proposed Approach
   - Progression
   - Incremental SDA

4 Experiments and Analysis
   - Protocol
   - Performance

5 Conclusion and Future work
**Discriminant Analysis Techniques**

![Discriminant Analysis Diagrams](image)

**Figure:** Illustration of where Linear Discriminant Analysis (left), Quadratic Discriminant Analysis (middle), and Subclass Discriminant Analysis (right) works.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Between-class scatter ($S_B$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDA</td>
<td>$\sum_{i=1}^{c} p_i (\mu_i - \mu)(\mu_i - \mu)^T$</td>
</tr>
<tr>
<td>QDA</td>
<td>$\sum_{i=1}^{c} p_i (\mu_i - \mu)(\mu_i - \mu)^T$</td>
</tr>
<tr>
<td>SDA</td>
<td>$\sum_{i=1}^{c-1} \sum_{j=1}^{H_i} \sum_{k=i+1}^{c} \sum_{l=1}^{H_k} p_{ij} p_{kl} (\mu_{ij} - \mu_{kl})(\mu_{ij} - \mu_{kl})^T$ (Encodes distance between sub-classes)</td>
</tr>
</tbody>
</table>

In SDA, the between-subclass scatter $S_B$ encodes distance between subclasses belonging to different classes.

The projection direction $w$ satisfies the Fisher criterion and ensures that subclasses (of different classes) are apart.
Subclass Discriminant Analysis

- In SDA, the between-subclass scatter $S_B$ encodes distance between subclasses belonging to different classes.
- The projection direction $w$ satisfies the Fisher criterion and ensures that subclasses (of different classes) are apart.

No built-in support for including new samples.
Requirement of Incremental Techniques

- In real world scenario, training samples are generally available in multiple smaller batches.
- The problem is to learn from **incremental** samples.
In real world scenario, training samples are generally available in multiple smaller batches.

The problem is to learn from **incremental** samples.
Requirement of Incremental Techniques

- In real world scenario, training samples are generally available in multiple smaller batches.
- The problem is to learn from **incremental** samples.
Outline

1 Motivation
   • Discriminant Analysis Techniques
   • Requirement of Incremental Techniques

2 Literature
   • Incremental PCA
   • Extension of IPCA to ILDA

3 Proposed Approach
   • Progression
   • Incremental SDA

4 Experiments and Analysis
   • Protocol
   • Performance

5 Conclusion and Future work
Related Incremental Approaches

- Hall et al. proposed incremental PCA by merging two eigenspaces.
- Kim et al. proposed incremental LDA by providing an approach to update discriminant components depending on merged eigenspaces of between and total scatter matrices.

For $i \in \{1, 2\}$ ($i = 1$ for existing batch, $i = 2$ for incremental batch),

- $N_i$ : number of samples in the $i^{th}$ batch
- $\mu_i$ : mean of the $i^{th}$ batch
- $C_i$ : covariance matrix of the $i^{th}$ batch

How to find the merged covariance $C_m$ matrix and new mean $\mu_m$?
**Incremental PCA - Approximate solution**

- $Ev_i$ : set of (selected first few) eigenvectors of covariance matrix $C_i$
- $\Lambda_i$ : diagonal matrix containing eigenvalues corresponding to the eigenvectors in $Ev_i$,
- $\{\mu_i, N_i, Ev_i, \Lambda_i\}$ ($i \in \{1, 2\}$) is eigenmodel of $C_i$

Find eigenmodel $\{\mu_m, N_m, Ev_m, \Lambda_m\}$ of merged covariance $C_m$.

\[
N_m = (N_1 + N_2)
\]
\[
\mu_m = (N_1\mu_1 + N_2\mu_2)/(N_m)
\]
\[
\Phi = QRDecomposition([Ev_1, Ev_2, (\mu_1 - \mu_2)])
\]
\[
\hat{P} \approx \Phi^T C_m \Phi
\]
\[
R = EigenVectors(\Phi^T C_m \Phi) \approx EigenVectors(\hat{P})
\]
\[
\Lambda_m = EigenValues(\Phi^T C_m \Phi) \approx EigenValues(\hat{P})
\]
\[
Ev_m = \Phi R
\]

- $\Phi^T C_m \Phi$ is calculated using both eigenmodels, without computing $C_m$
Incremental LDA

- Merged eigenmodels of existing batch $S_{B1}$ and incremental batch $S_{B2}$ to get eigenmodel of $S_{Bm}$.
- Similarly get merged eigenmodels of $S_{Tm}$ from $S_{T1}$ and $S_{T2}$ using Hall et al.
- Compute updated discriminative components $U$ using the merged eigenmodels.

Find updated discriminant components $U$ using eigenmodels $M_{Bi}$ and $M_{Wi}$ of scatter matrices $S_{Bi}$ and $S_{Wi}$ respectively. (i=1 for existing batch, i=2 for incremental batch).

$$S_{B2} = \sum_{i=1}^{c} p_i (\mu_i - \mu)(\mu_i - \mu)^T \quad \% \text{Between class scatter}$$

$$M_{Bm} = \text{merge}(M_{B1}, M_{B2}) \quad \% \text{by merging eigenspaces.}$$

$$M_{Tm} = \text{merge}(M_{T1}, M_{T2}) \quad \% \text{by merging eigenspaces.}$$

$$U = \text{getDiscComponents}(M_{Bm}, M_{Tm})$$
Incremental LDA

- Merged eigenmodels of existing batch $S_{B1}$ and incremental batch $S_{B2}$ to get eigenmodel of $S_{Bm}$.
- Similarly get merged eigenmodels of $S_{Tm}$ from $S_{T1}$ and $S_{T2}$ using Hall et al.
- Compute updated discriminative components $U$ using the merged eigenmodels.

Find updated discriminant components $U$ using eigenmodels $M_{Bi}$ and $M_{Wi}$ of scatter matrices $S_{Bi}$ and $S_{Wi}$ respectively. (i=1 for existing batch, i=2 for incremental batch).

$$S_{B2} = \sum_{i=1}^{c} p_i (\mu_i - \mu)(\mu_i - \mu)^T \quad \% \text{Between class scatter}$$

$$M_{Bm} = \text{merge}(M_{B1}, M_{B2}) \quad \% \text{by merging eigenspaces.}$$

$$M_{Tm} = \text{merge}(M_{T1}, M_{T2}) \quad \% \text{by merging eigenspaces.}$$

$$U = \text{getDiscComponents}(M_{Bm}, M_{Tm})$$
Merged eigenmodels of existing batch $S_{B1}$ and incremental batch $S_{B2}$ to get eigenmodel of $S_{Bm}$.

Similarly get merged eigenmodels of $S_{Tm}$ from $S_{T1}$ and $S_{T2}$ using Hall et al.

Compute updated discriminative components $U$ using the merged eigenmodels.

$S_{B2} = \sum_{i=1}^{c} p_i (\mu_i - \mu)(\mu_i - \mu)^T \% \text{ Between class scatter}$

$M_{Bm} = \text{merge}(M_{B1}, M_{B2}) \% \text{ by merging eigenspaces.}$

$M_{Tm} = \text{merge}(M_{T1}, M_{T2}) \% \text{ by merging eigenspaces.}$

$U = \text{getDiscComponents}(M_{Bm}, M_{Tm})$
Outline

1 Motivation
   - Discriminant Analysis Techniques
   - Requirement of Incremental Techniques

2 Literature
   - Incremental PCA
   - Extension of IPCA to ILDA

3 Proposed Approach
   - Progression
   - Incremental SDA

4 Experiments and Analysis
   - Protocol
   - Performance

5 Conclusion and Future work
IPCA proposed way to merge models of covariance matrices.

ILDA utilized it to merge between-class and total scatter matrices. Proposed technique to update discriminating components.

ISDA extends idea of ILDA, to incorporate between-sub-class scatter matrix.
Scatter Matrix for SDA

\[ S_{B2} = \sum_{i=1}^{c-1} \sum_{j=1}^{H_i} \sum_{k=i+1}^{c} \sum_{l=1}^{H_k} p_{ij} p_{kl} (\mu_{ij} - \mu_{kl})(\mu_{ij} - \mu_{kl})^T \]

- Sub-class labels of samples from the incremental batch are unknown.

Proposal: Find Sub-class labels using Nearest Neighbor. Using sub-class labels find sub-class means \( \mu_{ij} \).
  - Easy, simple and good to start with.
  - Does not incorporate varying number of sub-classes in run time.
Outline of the Proposed Incremental SDA

1. Find sub-class labels.
2. Compute between-subclass scatter matrix \( S_{B2} \) and total scatter matrix \( S_{M2} \) of incremental batch.
3. Find eigenmodels of scatter matrices. (Hall et al.)
4. Compute discriminative components. (Kim et al.)
The Proposed Incremental SDA

Find updated discriminant components $U$ using eigenmodels $M_{Bi}$ and $M_{Wi}$ of scatter matrices $S_{Bi}$ and $S_{Wi}$ respectively. (i=1 for existing batch, i=2 for incremental batch).

% Between sub-class scatter

$S_{B2} = \sum_{i=1}^{c-1} \sum_{j=1}^{H_i} \sum_{k=i+1}^{c} \sum_{l=1}^{H_k} p_{ij}p_{kl}(\mu_{ij} - \mu_{kl})(\mu_{ij} - \mu_{kl})^T$
The Proposed Incremental SDA

Find updated discriminant components $U$ using eigenmodels $M_{Bi}$ and $M_{Wi}$ of scatter matrices $S_{Bi}$ and $S_{Wi}$ respectively. ($i=1$ for existing batch, $i=2$ for incremental batch).

\[
S_{B2} = \sum_{i=1}^{c-1} \sum_{j=1}^{H_i} \sum_{k=i+1}^{c} \sum_{l=1}^{H_k} p_{ij}p_{kl}(\mu_{ij} - \mu_{kl})(\mu_{ij} - \mu_{kl})^T
\]

% Between sub-class scatter

% Merge eigenspaces (Hall et al.)

$M_{Bm} = \text{merge}(M_{B1}, M_{B2})$

$M_{Tm} = \text{merge}(M_{T1}, M_{T2})$
The Proposed Incremental SDA

Find updated discriminant components $U$ using eigenmodels $M_{Bi}$ and $M_{Wi}$ of scatter matrices $S_{Bi}$ and $S_{Wi}$ respectively. (i=1 for existing batch, i=2 for incremental batch).

% Between sub-class scatter
$$S_{B2} = \sum_{i=1}^{c-1} \sum_{j=1}^{H_i} \sum_{k=i+1}^{c} \sum_{l=1}^{H_k} p_{ij}p_{kl}(\mu_{ij} - \mu_{kl})(\mu_{ij} - \mu_{kl})^T$$

% Merge eigenspaces (Hall et al.)
$$M_{Bm} = \text{merge}(M_{B1}, M_{B2})$$
$$M_{Tm} = \text{merge}(M_{T1}, M_{T2})$$

% Compute discriminant components (Kim et al.)
$$Z = E_{Tm}^{-\frac{1}{2}}$$
$$\Omega = \text{QRDecomposition}(Z^T E_{Bm} V_{Bm})$$
$$R_D = \text{EigenVectors}(\Omega^T Z^T E_{Bm} V_{Bm} E_{Bm}^T Z \Omega)$$
$$U = Z \Omega R_D \text{% Updated discriminative components}$$
1 Motivation
   - Discriminant Analysis Techniques
   - Requirement of Incremental Techniques

2 Literature
   - Incremental PCA
   - Extension of IPCA to ILDA

3 Proposed Approach
   - Progression
   - Incremental SDA

4 Experiments and Analysis
   - Protocol
   - Performance

5 Conclusion and Future work
Protocol

- **Database**: AR face database
- **Face detection**: Viola-Jones face detector
- **Subjects**: 119 (face detection failed for other subjects)
- **Images**: 3094 (\(= 119 \times 26\))
- **Training : Testing** = 50% : 50% (non-overlapping)
- **Seen training**: (training set constitutes the gallery set)

<table>
<thead>
<tr>
<th>Batch</th>
<th>Images per Subject</th>
<th>Total Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch1</td>
<td>9</td>
<td>1071</td>
</tr>
<tr>
<td>Batch2</td>
<td>2</td>
<td>238</td>
</tr>
<tr>
<td>Batch3</td>
<td>2</td>
<td>238</td>
</tr>
<tr>
<td>Total Training</td>
<td>13</td>
<td>1547</td>
</tr>
</tbody>
</table>

Performance (Time)

<table>
<thead>
<tr>
<th></th>
<th>PCA</th>
<th>CCIPCA</th>
<th>LDA</th>
<th>ILDA</th>
<th>SDA</th>
<th>ISDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial batch (batch I)</td>
<td>69.6</td>
<td>18</td>
<td>15.2</td>
<td>11.4</td>
<td>6167.7</td>
<td>6712</td>
</tr>
<tr>
<td>Batch II</td>
<td>83.3</td>
<td>22.3</td>
<td>19.5</td>
<td>13.5</td>
<td>10610</td>
<td>22.2</td>
</tr>
<tr>
<td>Batch III</td>
<td>99.8</td>
<td>26.3</td>
<td>20.2</td>
<td>25.4</td>
<td>17494</td>
<td>24.6</td>
</tr>
</tbody>
</table>

**Table:** Time taken (in seconds) by each of the approaches for initial training (Batch I) and incremental training (Batch II and III)

- Time expensive behavior of SDA contributes to high time complexity of estimating the numbers of subclasses.

---

**Performance (Accuracy)**

![Graphs showing performance accuracy](image)

- **(a) Batch I (Initial)**
- **(b) Batch II (Update)**
- **(c) Batch III (Convergence)**

**Figure:** CMC plots of the proposed ISDA and performance comparison with PCA, CCIPCA, LDA, ILDA, and SDA
Figure: Rank 5 identification accuracy averaged across 5 cross validations.
**Table:** The co-occurrence of correct classifications (✓) and/or misclassifications (✗) between SDA and ISDA. It turns out that for only \( \frac{87+50}{1547} \times 100\% = 8.85\% \) of the times, the decisions taken by SDA and ISDA differ. Table shows the confusion matrix for batch III at rank 5.

<table>
<thead>
<tr>
<th>Confusion matrix @ Rank 5</th>
<th>SDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ISDA</td>
<td>1091, 50</td>
</tr>
<tr>
<td>✗</td>
<td>87, 319</td>
</tr>
</tbody>
</table>
Outline

1. Motivation
   - Discriminant Analysis Techniques
   - Requirement of Incremental Techniques

2. Literature
   - Incremental PCA
   - Extension of IPCA to ILDA

3. Proposed Approach
   - Progression
   - Incremental SDA

4. Experiments and Analysis
   - Protocol
   - Performance

5. Conclusion and Future work
Formulation of ISDA is proposed, and evaluated in the scenario of face recognition.

ISDA achieves almost same identification accuracy, as SDA, is achieved; that too,  
- without recomputing scatter matrices from scratch  
- with lower training time complexity.
Future work

- Ability to add new classes in incremental manner.
- Evaluating performance on other pattern classification problems.
Thank You!!
Questions are always welcomed !!