Discrimination of Dyed Cotton Fibers Based on UV-visible Microspectrophotometry and Multivariate Statistical Analysis

Elisa Liszewski, Cheryl Szkudlarek and John Goodpaster
Department of Chemistry and Chemical Biology
Forensic and Investigative Sciences Program
Indiana University Purdue University Indianapolis (IUPUI)
Indianapolis, IN  46202
Objectives

1. Acquire a collection of dyed cotton fibers
2. Acquire spectra from all samples using UV-visible microspectrophotometers at IUPUI and the Indiana State Police
3. Carry out multivariate statistical analysis on the resultant spectra
4. Form a dye database
Questions We Can Answer

- How many classes of spectra can be reliably discerned in a population of dyed cotton fibers that have been analyzed by microspectrophotometry?
- What general features of the spectra represent the defined groups so that an unknown spectrum could be tentatively classified?
- What regions of the spectra are most important for discriminating these groups and therefore are the most reliable regions to inspect when comparing samples?
- To what extent can an unknown sample be correctly and quantitatively assigned to its member class?
- Can spectra obtained on one instrument serve as a database for a different instrument?
- If not, does the discrimination of spectra differ significantly depending on the instrument used?
- Does the calculation of first-derivative spectra or the use of chromaticity coordinates offer any advantages when comparing samples?
- Ultimately, to what extent do real fiber samples exhibit heterogeneity and what effect does this have on declaring a known and questioned sample to be indistinguishable with respect to their absorption characteristics?
Fibers as Trace Evidence

• Why are Fibers Important?
  – Widespread in the environment
  – Transferred easily
  – Many classifications and subtypes
  – Physically and chemically differentiable

• Key Characteristics
  – Morphology (cross-section)
  – Bulk Composition
  – Color *

*This is the main discriminating feature for cotton fibers
Fiber Dyes

• What is a Dye?
  – A colored substance that is able to absorb and reflect certain visible wavelengths of light

• Thousands of dyes produced worldwide

• Classification of Dyes:
  – By method of application
  – By chemical class
  – By type of fiber to which they are applied
Classification by Application Method

- **Acid**
  - Neutral or Acidic conditions forms ionic bonds

- **Azoic**
  - Bonding between diazo and coupling component

- **Basic**
  - Acidic conditions forms ionic bonds

- **Direct**
  - Directly incorporated into cellulosic fiber with the presence of heat and an electrolyte

- **Dispersive**
  - Directly incorporated into synthetic fibers

- **Reactive**
  - Form covalent bonds with the functional groups

- **Sulfur**
  - Require reducing agent to make dye soluble and then oxidize within the fiber to become insoluble

- **Vat**
  - Require reducing agent to make dye soluble and then oxidize within the fiber to become insoluble
The Dyes

<table>
<thead>
<tr>
<th>Label</th>
<th>Dye</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Direct Red C-380</td>
</tr>
<tr>
<td>B</td>
<td>Reactive Red 120</td>
</tr>
<tr>
<td>C</td>
<td>Reactive Red 123</td>
</tr>
<tr>
<td>D</td>
<td>Reactive Red 195</td>
</tr>
<tr>
<td>E</td>
<td>Reactive Red 2</td>
</tr>
<tr>
<td>F</td>
<td>Reactive Red 228</td>
</tr>
</tbody>
</table>

*Provided by Testfabsics*
Instrumental Analysis of Dyed Textile Fibers

- Relies upon principles of:
  - microscopy
  - spectroscopy
  - chromatography
  - mass spectrometry

Sample Preparation

- 6 Dyed exemplars
- 10 fibers/dye
- Mounted in glycerin on glass slides
- CRAIC QDI 2000 MSP
- 35x magnification
- MSP calibrated prior to analysis
- 10 scans/fiber
Data Analysis

• Spectra truncated to 350-800 nm range
• Background subtracted
• Normalized

• Chemometric techniques run:
  – Agglomerative Hierarchical Clustering (AHC)
  – Principal Components Analysis (PCA)
  – Discriminant Analysis (DA)
  – Analysis of Variance (ANOVA)
Representative Spectra

- Direct Red C-380
- Reactive Red 120
- Reactive Red 123
- Reactive Red 195
- Reactive Red 2
- Reactive Red 228

WAVELENGTH (nm)
Hierarchial Cluster Analysis (AHC)

• Purpose:
  – Present data in a way that emphasize groupings
  – Detect outliers

• Procedure:
  – Calculate and Compare distances between samples
    • Small distances indicate similarity

• Result
  – Dendrogram
    • Display results according to dissimilarity
    • Truncation line
      – Groupings to the right significant
Dendrogram

Reactive Red 120, Reactive Red 2, Reactive Red 228

Reactive Red 195

Direct Red C-380, Reactive Red 120, Reactive Red 2

Reactive Red 123
# Dendrogram Summary

<table>
<thead>
<tr>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Red C-380</td>
<td>Reactive Red 120</td>
<td>Reactive Red 123</td>
<td>Reactive Red 195</td>
</tr>
<tr>
<td>Reactive Red 120</td>
<td>Reactive Red 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reactive Red 2</td>
<td>Reactive Red 228</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Unique Fiber Detection

FIBER E

NORMALIZED ABSORBANCE vs WAVELENGTH (nm)
Principle Component Analysis (PCA)

• Purpose:
  – Reduce dimensionality of data
  – Detect patterns and outliers

• Procedure:
  – Form principle components
    • First pc captures the most variance

• Result:
  – Observation Plot
    • First two principle components
  – Factor Loading Plot
    • Contributions of the variables
Observation (axes F1 and F2: 67.51 %)
PCA Observations with 3 factor varimax rotation
Factor Loading

WAVELENGTH (nm)

D1
D2
Most Informative Regions of Spectra

PC1 Positive Correlation: 350-387 nm
PC1 Negative Correlation: 502-513 nm
PC2 Positive Correlation: 402-484 nm
PC2 Negative Correlation: 528-565 nm
Linear Discriminant Analysis (DA)

• Purpose:
  – Predict group membership
  – Detect patterns

• Procedure:
  – Form canonical variates
  – Assign to class with highest probability

• Result:
  – Observations Plot
    • First two canonical variates
  – Confusion Matrix
    • Summary of the reclassification
DA Results

- The accuracy of DA can be measured by leave-one-out cross validation.
- Groups listed with 100% accuracy had no errors in re-classification.
- Groups with low accuracy are easily confused with other dyes.

<table>
<thead>
<tr>
<th>Fiber</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Red C-380</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 120</td>
<td>69.00</td>
</tr>
<tr>
<td>Reactive Red 123</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 195</td>
<td>98.00</td>
</tr>
<tr>
<td>Reactive Red 2</td>
<td>59.00</td>
</tr>
<tr>
<td>Reactive Red 228</td>
<td>85.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>85.26</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fiber</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Red C-380</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 120</td>
<td>74.23</td>
</tr>
<tr>
<td>Reactive Red 123</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 195</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 2</td>
<td>87.00</td>
</tr>
<tr>
<td>Reactive Red 228</td>
<td>91.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>92.13</strong></td>
</tr>
</tbody>
</table>
Analysis of Variance (ANOVA)

• **Purpose:**
  – Find parts of the spectra for maximum discrimination

• **Procedure:**
  – Assign F-values
    • \( F = \frac{\text{Between Group Variance}}{\text{Within Group Variance}} \)

• **Result:**
  – F-values
  – Fisher Ratio Plot
    • High value indicates difference
Univariate Fisher Ratios

WAVELENGTH (nm)
Interestingly, the most discriminating regions do not include the $\lambda_{\text{max}}$ – rather they lie on the rising and falling edges.
The Results

EXTERNAL VALIDATION
Observations (axes F1 and F2: 99.08 %)
## DA Results

<table>
<thead>
<tr>
<th>Fiber</th>
<th>% Correct</th>
<th>Fiber</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Red C-380_1</td>
<td>100.00</td>
<td>Direct Red C-380_1</td>
<td>100.00</td>
</tr>
<tr>
<td>Direct Red C-380_2</td>
<td>100.00</td>
<td>Direct Red C-380_2</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 120_1</td>
<td>90.00</td>
<td>Reactive Red 120_1</td>
<td>70.00</td>
</tr>
<tr>
<td>Reactive Red 120_2</td>
<td>90.00</td>
<td>Reactive Red 120_2</td>
<td>90.00</td>
</tr>
<tr>
<td>Reactive Red 123_1</td>
<td>100.00</td>
<td>Reactive Red 123_1</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 123_2</td>
<td>100.00</td>
<td>Reactive Red 123_2</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 195_1</td>
<td>100.00</td>
<td>Reactive Red 195_1</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 195_2</td>
<td>90.00</td>
<td>Reactive Red 195_2</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 2_1</td>
<td>10.00</td>
<td>Reactive Red 2_1</td>
<td>40.00</td>
</tr>
<tr>
<td>Reactive Red 2_2</td>
<td>30.00</td>
<td>Reactive Red 2_2</td>
<td>10.00</td>
</tr>
<tr>
<td>Reactive Red 228_1</td>
<td>90.00</td>
<td>Reactive Red 228_1</td>
<td>70.00</td>
</tr>
<tr>
<td>Reactive Red 228_2</td>
<td>90.00</td>
<td>Reactive Red 228_2</td>
<td>90.00</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>82.50</strong></td>
<td><strong>TOTAL</strong></td>
<td><strong>80.83</strong></td>
</tr>
</tbody>
</table>

* 6 Class Assignment

* 60 Class Assignment
IUPUI versus ISP

INTER-LABORATORY STUDY
# The Dyes

<table>
<thead>
<tr>
<th>Label</th>
<th>Dye</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Direct Red C-380</td>
</tr>
<tr>
<td>B</td>
<td>Reactive Red 120</td>
</tr>
<tr>
<td>C</td>
<td>Reactive Red 123</td>
</tr>
<tr>
<td>D</td>
<td>Reactive Red 195</td>
</tr>
<tr>
<td>E</td>
<td>Reactive Red 2</td>
</tr>
<tr>
<td>F</td>
<td>Reactive Red 228</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Label</th>
<th>Dye</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>Direct Red 84</td>
</tr>
<tr>
<td>H</td>
<td>Reactive Red 180</td>
</tr>
<tr>
<td>I</td>
<td>Reactive Red 198</td>
</tr>
<tr>
<td>J</td>
<td>Reactive Red 239/241</td>
</tr>
<tr>
<td>K</td>
<td>Vat Red 10</td>
</tr>
<tr>
<td>L</td>
<td>Vat Red 15</td>
</tr>
</tbody>
</table>

*Provided by Testfabrics*

*Provided by Dr. Stephen Morgan from University of South Carolina*
<table>
<thead>
<tr>
<th>Dye</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Red C-380</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 120</td>
<td>67.00</td>
</tr>
<tr>
<td>Reactive Red 123</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 195</td>
<td>72.00</td>
</tr>
<tr>
<td>Reactive Red 2</td>
<td>50.00</td>
</tr>
<tr>
<td>Reactive Red 228</td>
<td>83.00</td>
</tr>
<tr>
<td>Direct Red 84</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 180</td>
<td>89.00</td>
</tr>
<tr>
<td>Reactive Red 198</td>
<td>92.00</td>
</tr>
<tr>
<td>Reactive Red 239/241</td>
<td>67.00</td>
</tr>
<tr>
<td>Vat Red 10</td>
<td>100.00</td>
</tr>
<tr>
<td>Vat Red 15</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>85.00</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dye</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Red C-380</td>
<td>92.00</td>
</tr>
<tr>
<td>Reactive Red 120</td>
<td>61.00</td>
</tr>
<tr>
<td>Reactive Red 123</td>
<td>100.00</td>
</tr>
<tr>
<td>Reactive Red 195</td>
<td>93.00</td>
</tr>
<tr>
<td>Reactive Red 2</td>
<td>39.00</td>
</tr>
<tr>
<td>Reactive Red 228</td>
<td>77.00</td>
</tr>
<tr>
<td>Direct Red 84</td>
<td>99.00</td>
</tr>
<tr>
<td>Reactive Red 180</td>
<td>94.00</td>
</tr>
<tr>
<td>Reactive Red 198</td>
<td>70.00</td>
</tr>
<tr>
<td>Reactive Red 239/241</td>
<td>52.00</td>
</tr>
<tr>
<td>Vat Red 10</td>
<td>99.00</td>
</tr>
<tr>
<td>Vat Red 15</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>81.25</strong></td>
</tr>
</tbody>
</table>
Conclusions

• Training Set:
  – 4 main classes
  – 67.51% total variance captured by first 2 PCs
  – Direct Red C-380, Reactive Red 123, and Reactive Red 195 were correctly classified 100%
  – Overall accuracy of classification was above 90%
  – 463-502 nm and 554-585 nm were the most discriminating regions
  – Some “uniqueness was seen within the 10 fibers of Reactive Red 2
More Conclusions

• External Validation:
  – The overall classification accuracy was above 80%

• Inter-laboratory Study:
  – Five dyes were readily distinguished using instruments at IUPUI and ISP:
    • Direct Red C-380, Reactive Red 123, Direct Red 84, Vat Red 10, and Vat Red 15
  – The remaining dyes were not as readily distinguished and potentially confused with one another – the degree of confusion varied between laboratories
  – Overall, consistency was shown between the two instruments
Acknowledgements

• Microanalysis Unit of the Indiana State Police Laboratory
• Dr. Stephen Morgan (University of South Carolina)
• Tom Klass (Testfabrics, Inc.)
• Midwest Forensics Resource Center (MFRC)