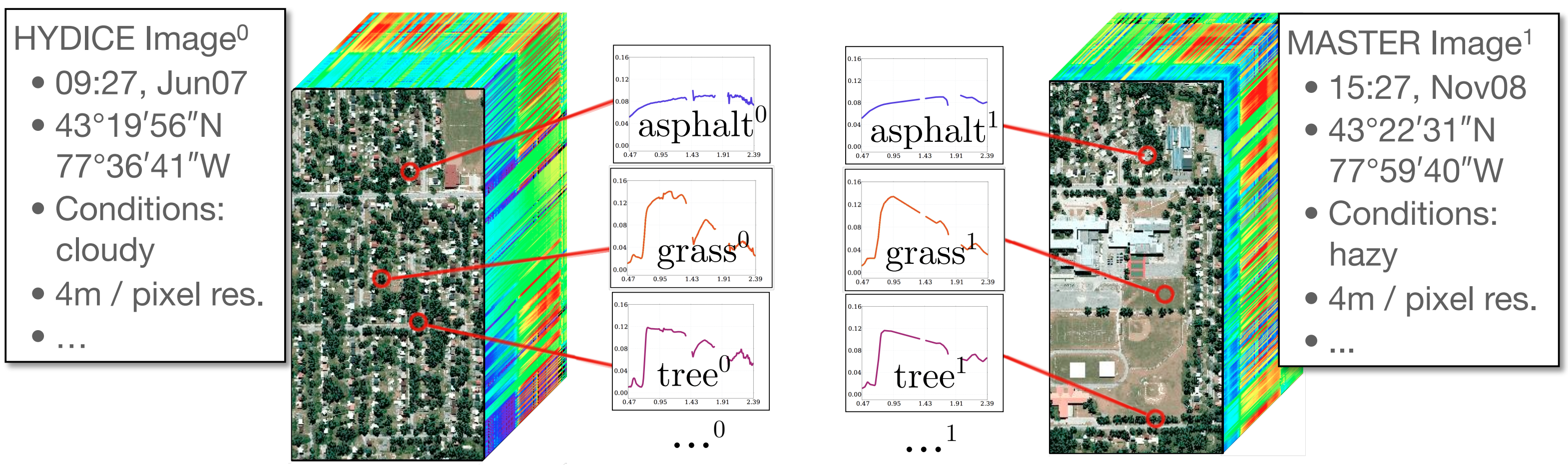
An Evaluation of Class Knowledge Transfer from Synthetic to Real Hyperspectral Imagery

Brian Bue[1], Erzsébet Merényi[2], Beáta Csathó[3]
[1] Electrical and Computer Engineering, Rice University, Houston, TX
[2] Statistics, Rice University, Houston, TX
[3] Geology, University at Buffalo, SUNY, Buffalo, NY

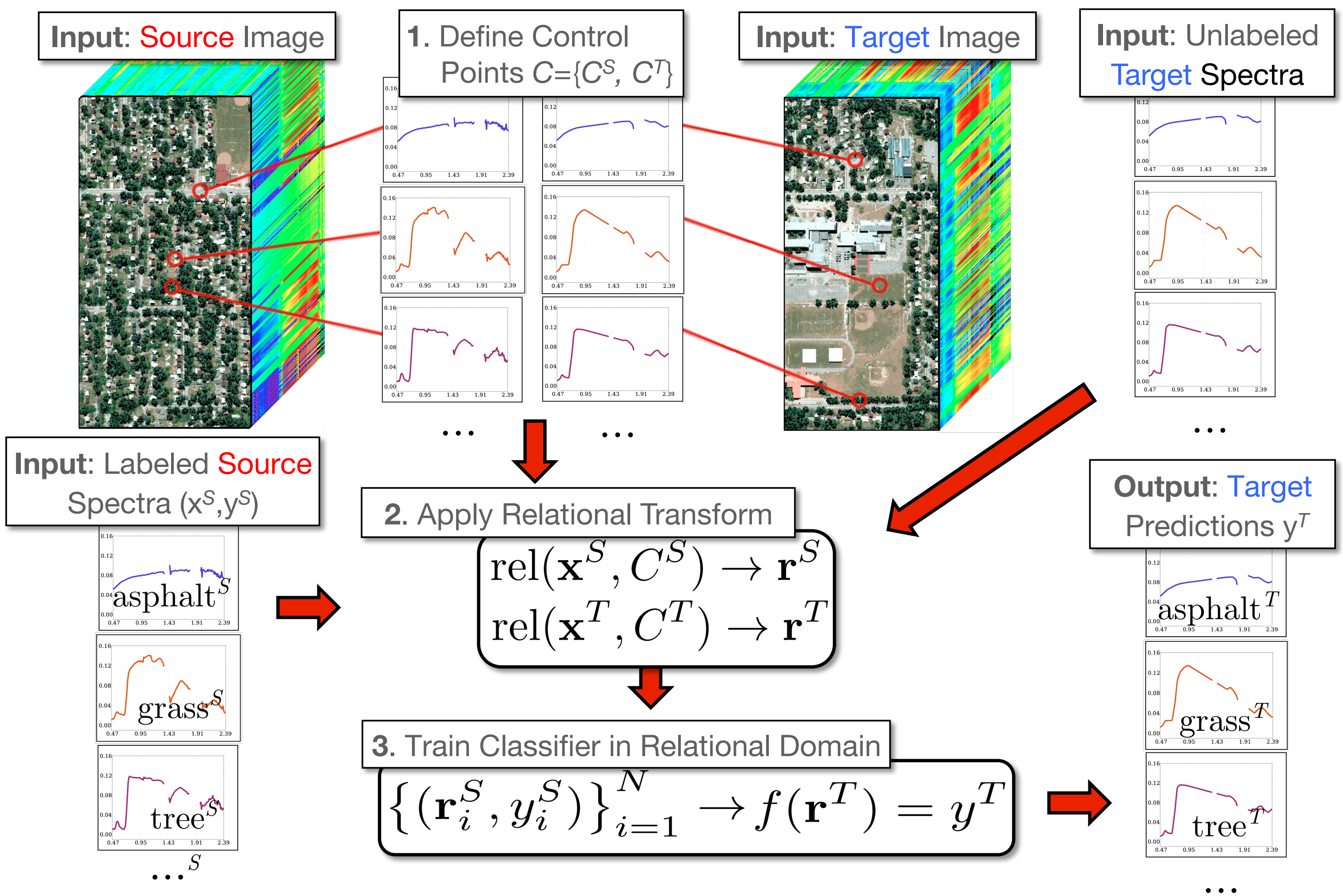
Transferring Class Knowledge Between Images

- Images captured under differing conditions (e.g., different sensors, atmospheric conditions, spatial locations, capture dates) commonly share similar materials



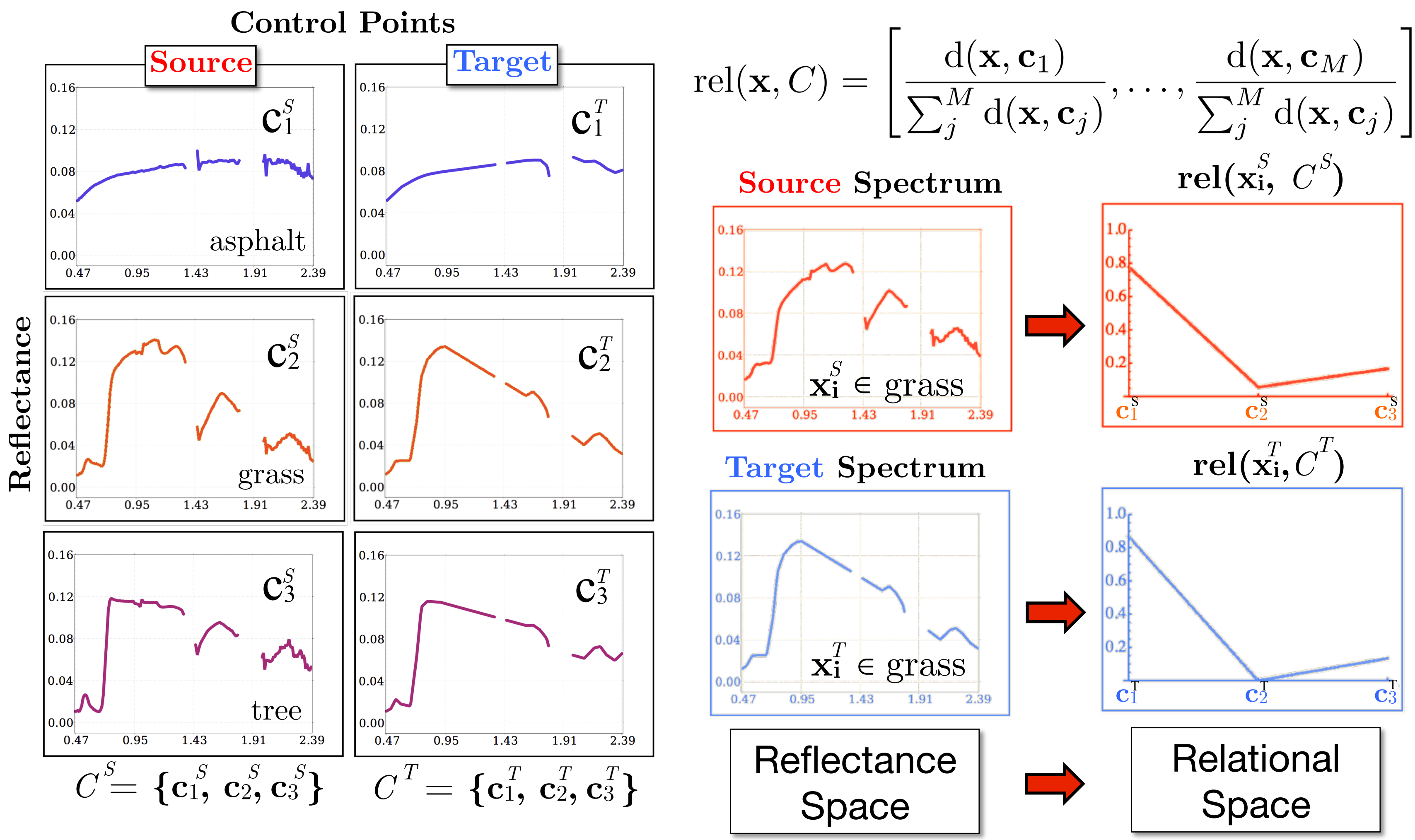
- Reconciling differences between images allows us to train a classifier using labeled spectra from one image (source) to label other, similar (target) imagery
- Reduces, expensive, tedious and error-prone manual labeling
- Facilitates class knowledge transfer from synthetic to real imagery

Relational Class Knowledge Transfer (RelTrans) [Bue et al., 2010]

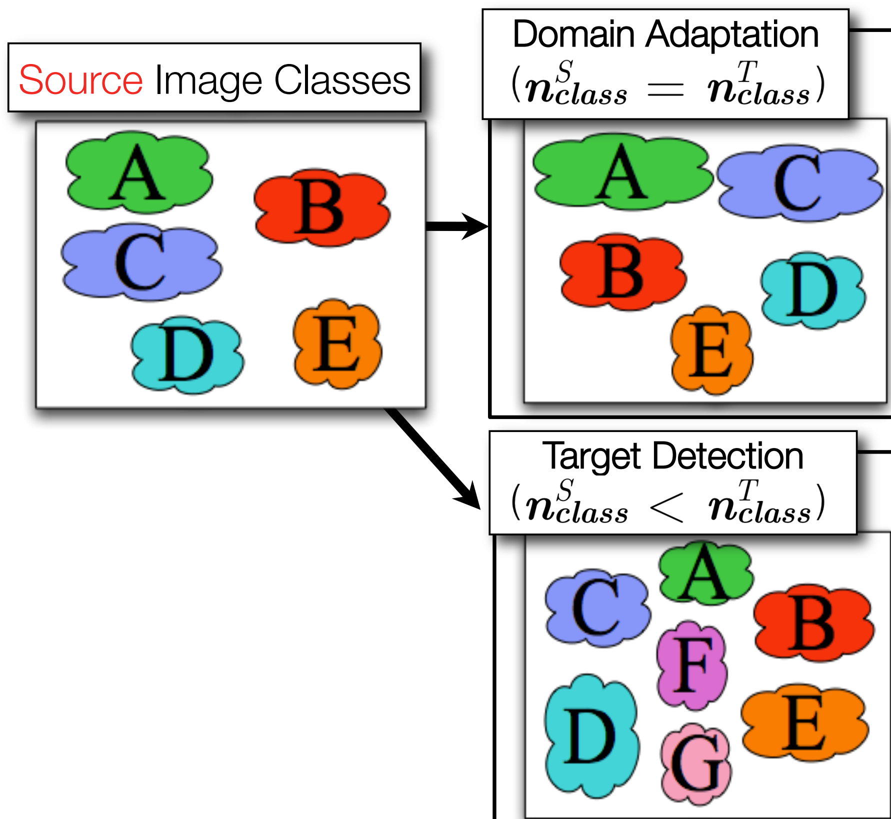


RelTrans Transform

$$\mathrm{rel}(\mathbf{x}, C) = \left[\frac{\mathrm{d}(\mathbf{x}, \mathbf{c}_1)}{\sum_j^M \mathrm{d}(\mathbf{x}, \mathbf{c}_j)}, \ldots, \frac{\mathrm{d}(\mathbf{x}, \mathbf{c}_M)}{\sum_j^M \mathrm{d}(\mathbf{x}, \mathbf{c}_j)}\right]$$



Domain Adaptation & Target Detection Scenarios



- All classes in source image present in target image
- At least 1 control point per source class
- Transfer accuracy = % of correctly labeled target pixels

- Some classes in target image not present in target image
- At least 1 control point per source class
- Transfer accuracy = % of correctly labeled target pixels belonging to source classes
    - Pixels incorrectly flagged as unknowns considered errors
- Accuracy reported on unflagged pixels only

Target Detection: RelThresh

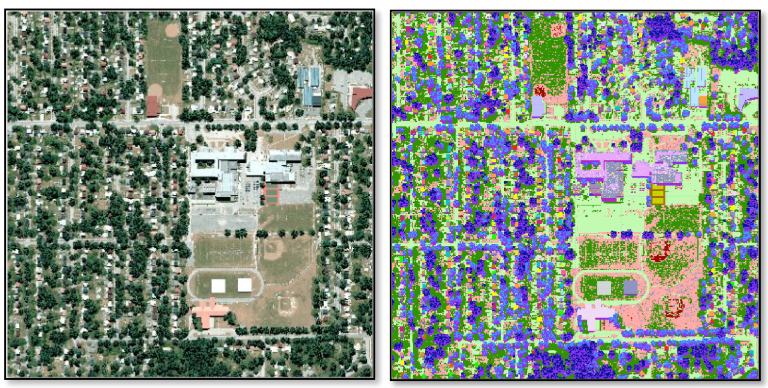
- Target imagery will often contain classes not present in source imagery
  - Need a mechanism to flag pixels as “unknowns”
- Class likelihoods = relsim distances to source and target control point class means:

$$\mathrm{relsim}(\mathbf{x}_i, \boldsymbol{\mu}_j, C) = 1 - \frac{k}{2}||\mathrm{rel}(\mathbf{x}_i, C) - \mathrm{rel}(\boldsymbol{\mu}_j, C)||_2$$

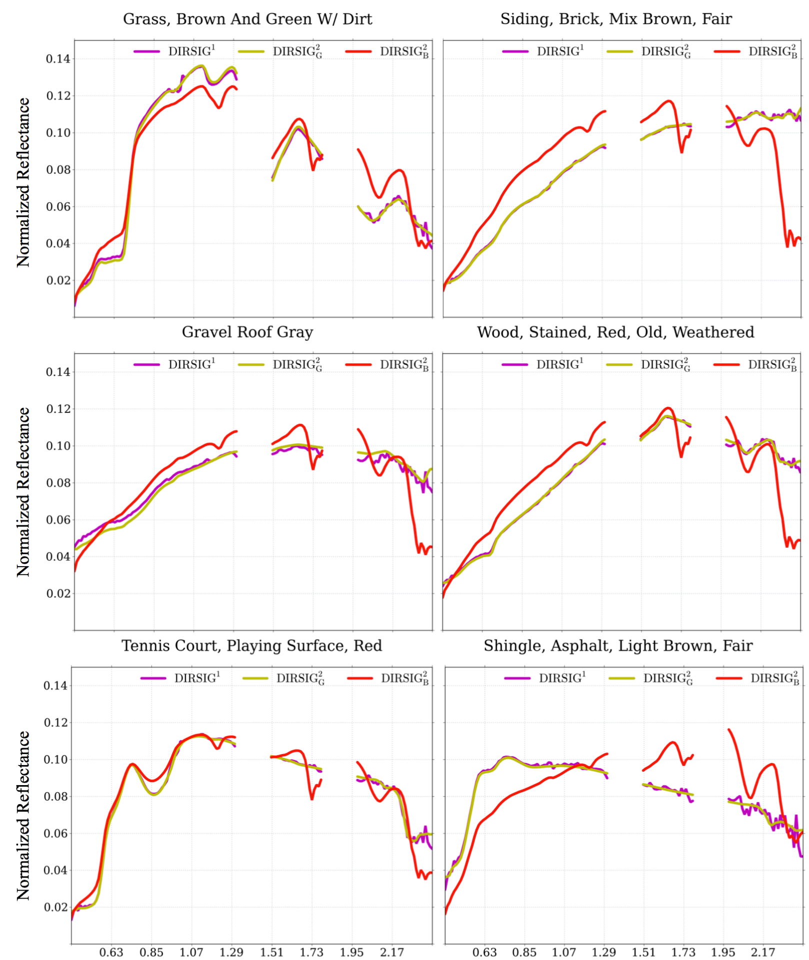
- RelThresh procedure: select T as the maximum relsim likelihood value which flags no control points as unknowns
- Advantages:
  - RelTrans feature space allows for single threshold for all classes
  - Simple to compute (linear scan of control point likelihood values)

Synthetic Transfer Assessment: Distorted vs. Clean Source Spectra

- Target image: DIRSIG[1]: 210 band, 4m/pixel DIRSIG [Schott et al 1999] synthetic HYDICE
- Source images:
  - DIRSIG[2]G: “cleaner” DIRSIG[1] image with reduced atmospheric effects and fewer shadow pixels
  - DIRSIG[2]B: poorly atmospherically calibrated version of DIRSIG[2]G image



DIRSIG[1] RGB Composite and Material Map



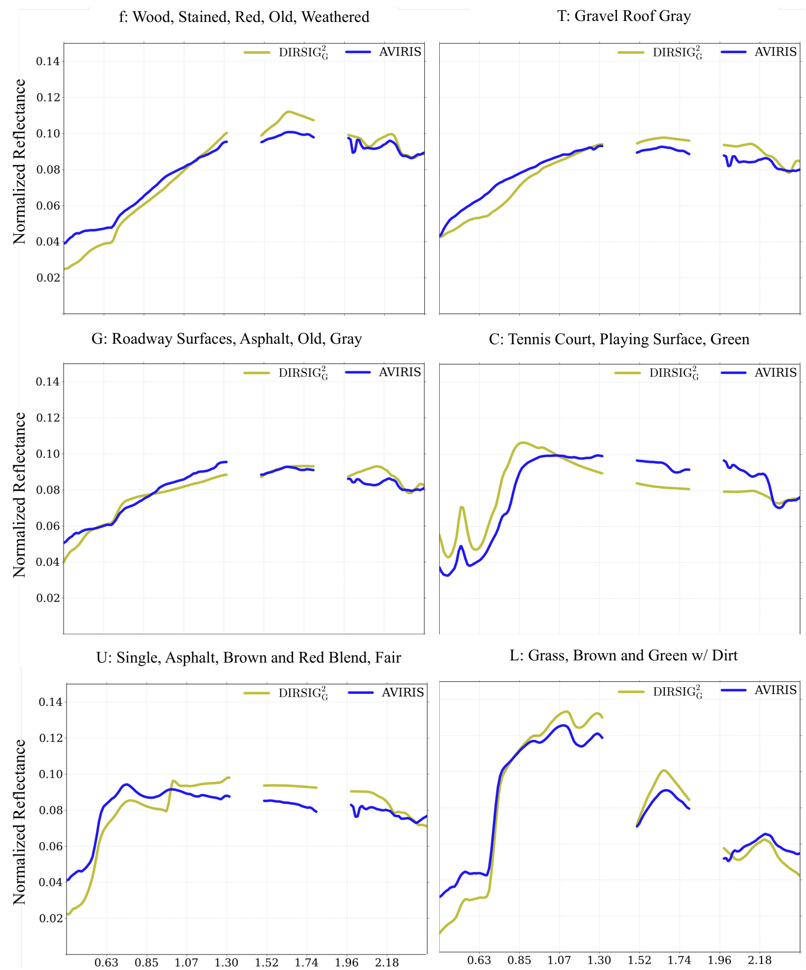
Mean signatures: DIRSIG[1] and DIRSIG[2] control points

Synthetic to Real Class Knowledge Transfer

- Target image: 224-band AVIRIS image of Ocean City, MD
  [Csathó et al., 1998, Merényi et al 2009]
    - Common materials to source and target images: road/sidewalk (asphalts, concrete), building materials (siding, rooftops), vegetation (grass, trees)
- Source image: DIRSIG^{2}G HYDICE image

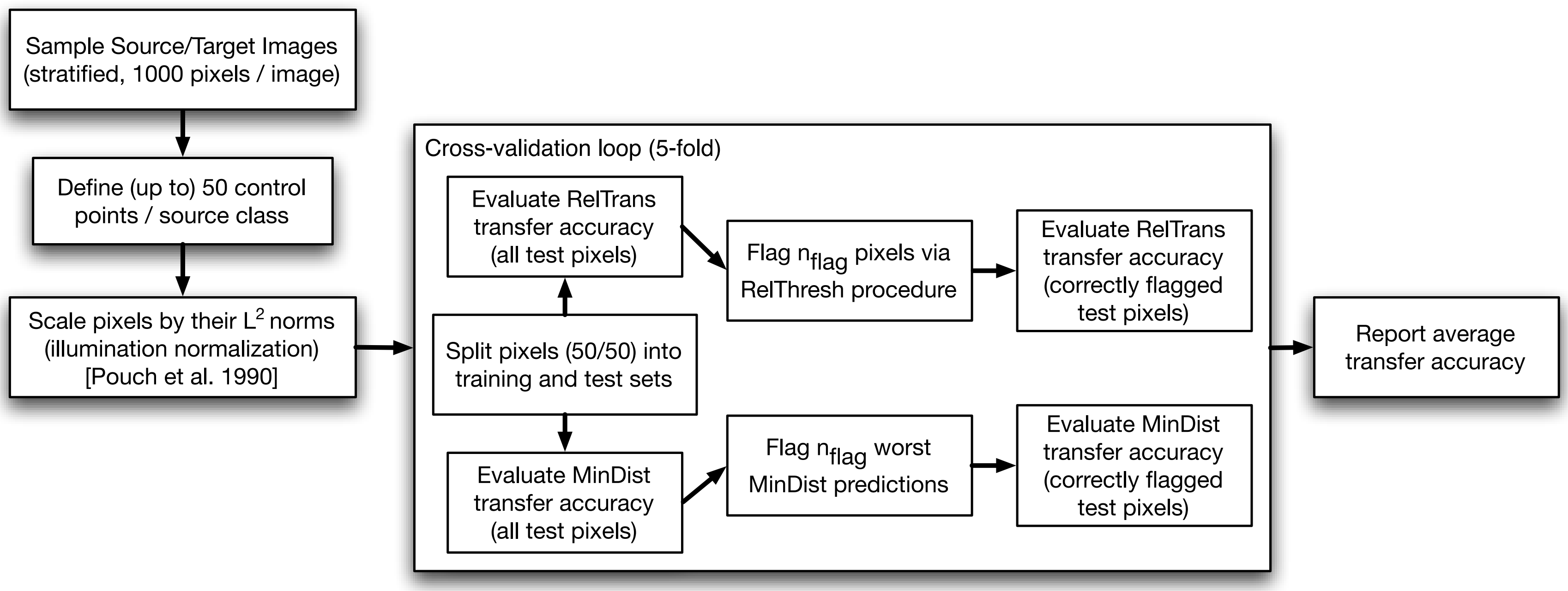
Ocean City, MD
AVRISLA Image

Segmentation Map

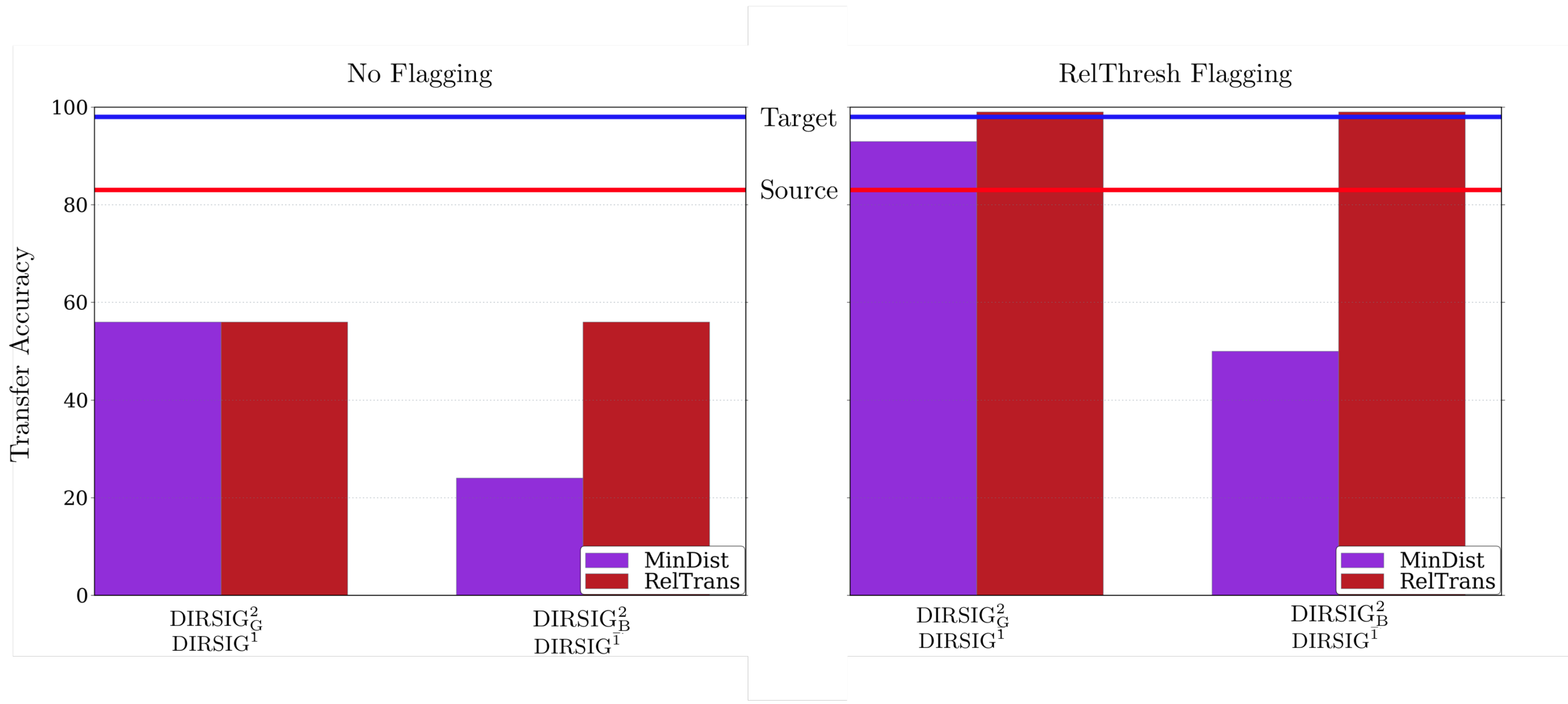


Mean signatures: DIRSIG^{2}G, AVIRIS control points

Evaluation Methodology



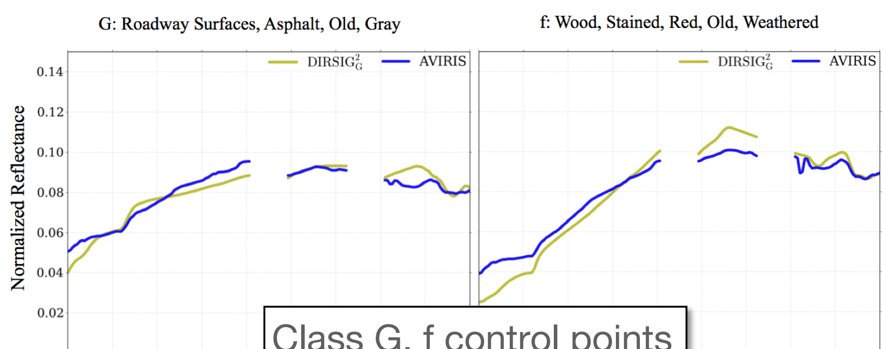
Synthetic Image Target Detection Results



- 6 common classes (561 samples) + 5 target-only classes (439 samples)
  - Max possible transfer accuracy w/o flagging = 56%
- All target-only pixels correctly flagged by RelThresh

Synthetic to Real Domain Adaptation Results

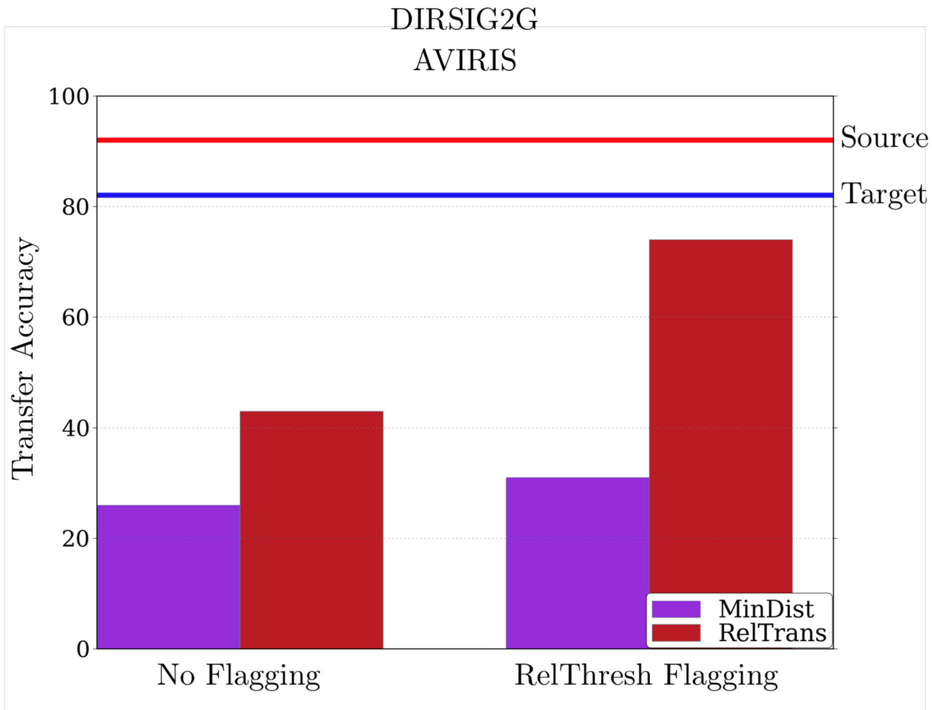
Segment Label: Material Class	MinDist	RelTrans
C: Tennis Court, Playing Surface, Green	7	100
G: Roadway Surfaces, Asphalt, Old, Gray	55	26
L: Grass, Brown and Green w/ Dirt	63	93
T: Gravel Roof Gray	0	100
U: Shingle, Asphalt, Brown and Red Blend, Fair	57	100
f: Wood, Stained, Red, Old, Weathered	28	83



Class G, f control points

- Highly correlated control points may reduce transfer accuracy
  - Particularly if correlated in only one of the source or target spaces
- Filtering/decorrelating control points could improve results

Synthetic to Real Target Detection Results



- 6 common classes (539 samples), 5 target-only classes (431 samples)
  - Maximum transfer accuracy w/o flagging: 54%
- RelTrans nearly achieves target classification accuracy (despite correlated control points)

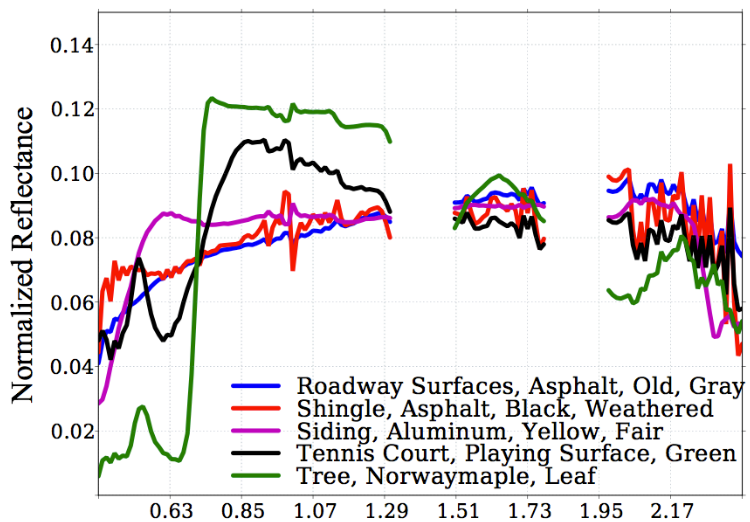
Conclusions / Future Work

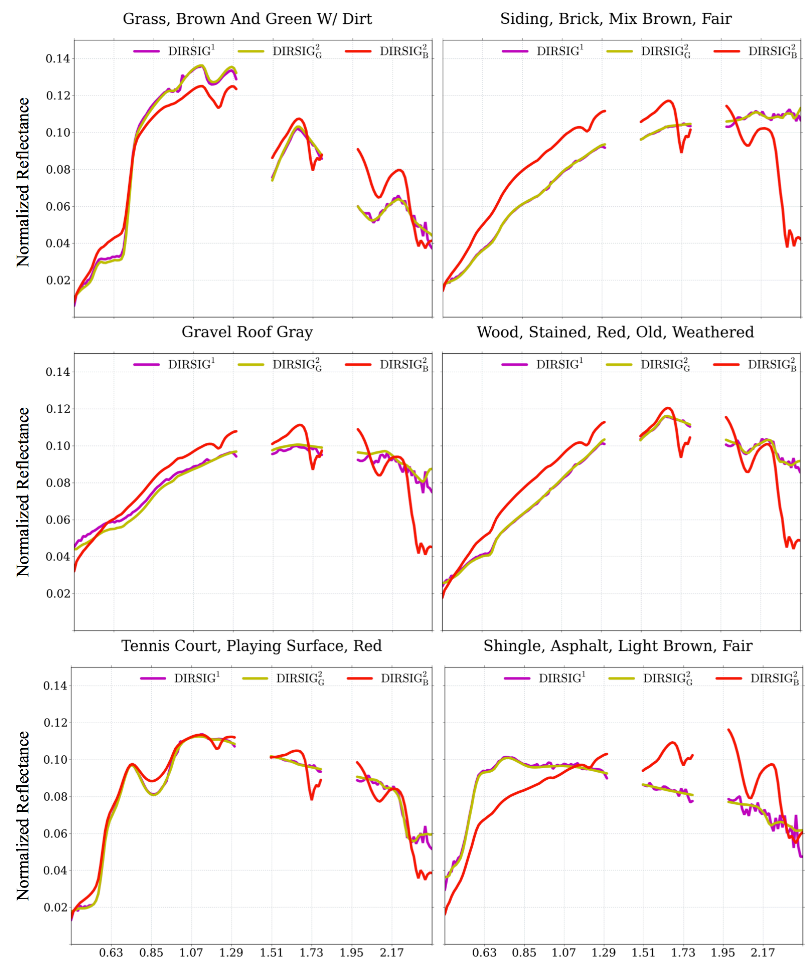
- When two images share some material classes, it is possible to train accurate classifiers using only one of the images (the source) as training data with a few source-to-target control points
- Control points provide useful information to distinguish between classes present only in the source or target images
- Classifying target pixels using the RelTrans transform performs best when control points are not highly correlated in source and target imagery
- Currently exploring methods to automatically detect source and target correspondence points
- Using adaptive spectral similarity measures in the RelTrans transform may improve knowledge transfer performance

References

- B Bue and E Merényi, “Using spatial correspondences for hyperspectral knowledge transfer: evaluation on synthetic data,” IEEE WHISPERS 2010.
- GW Pouch and DJ Campagna, “Hyperspherical direction cosine transformation for separation of spectral and illumination information in digital scanner data,” Photogrammetric Engineering and Remote Sensing, vol. 56, no. 4, pp. 475–479, 1990.
- B Csathó, WB Krabill, J Lucas, and T Schenk, “A multisensor data set of an urban and coastal scene,” International Archives of Photogrammetry and Remote Sensing, vol. XXXII (3/2), pp. 26–31, Jan 1998.
- E Merényi, K Tasdemir, and W Farrand, “Intelligent information extraction to aid science decision making in autonomous space exploration,” Proceedings of SPIE, vol. 6960, 2009.
- JR Schott, SD Brown, RV Raqueno, HN Gross, and G Robinson, “Advanced synthetic image generation models and their application to multi/hyperspectral algorithm development,” Proceedings of SPIE, vol. 3584, pp. 211, 1999.

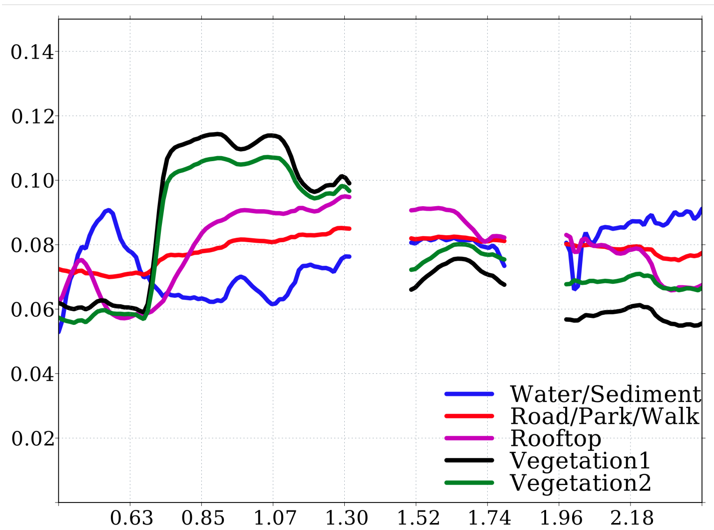
Target Detection: Synthetic Data



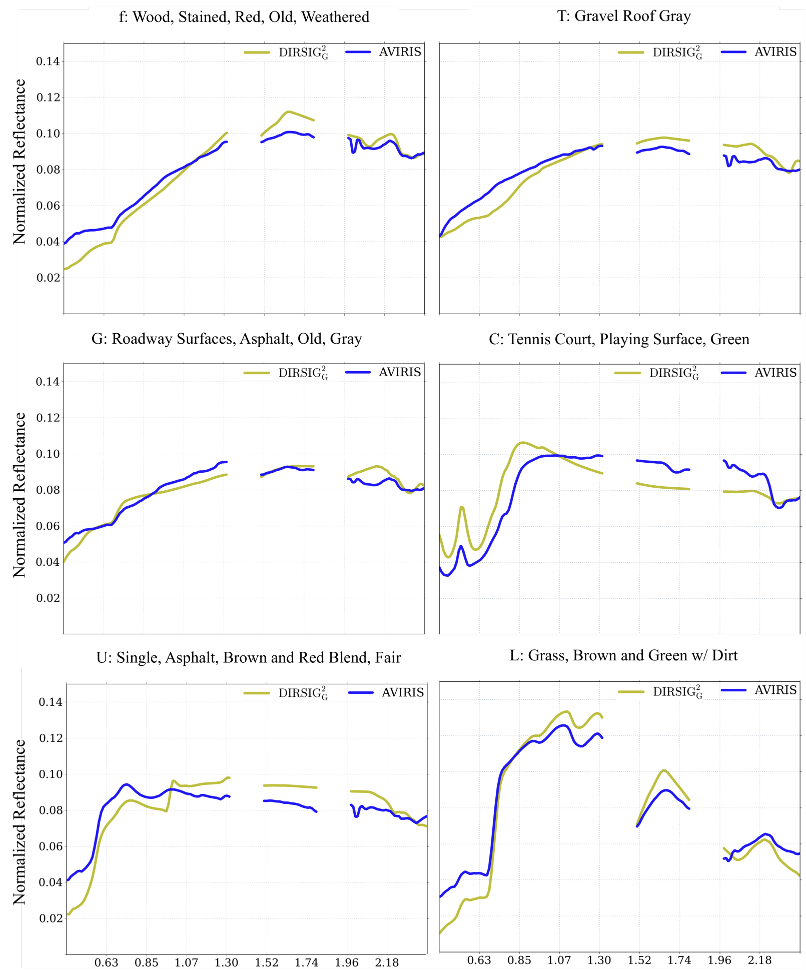


Target Detection	DIRSIG^2_B(83) vs. DIRSIG^1(98)
MinDist	24 / 50 [17]
RelTrans	56 / 99 [100]

Target Detection: Synthetic-to-Real Data



Target Detection	DIRSIG^2_G(92) vs. AVIRIS(82)
MinDist	26 / 31 [77]
RelTrans	43 / 74 [100]



RelTrans Algorithm

Algorithm 1 RelTrans

Input: $n_S \times d$ matrix of source pixels $\mathbf{S}$. $n_T \times d$ matrix of target pixels $\mathbf{T}$. Set of n_C correspondence points $C = \left\{(\mathbf{c}_1^S, \mathbf{c}_1^T, l_1^S), \ldots, (\mathbf{c}_{n_C}^S, \mathbf{c}_{n_C}^T, l_{n_C}^S)\right\}$. Similarity threshold τ.

Output: $n_T \times k_S$ class similarity matrix $\mathbf{\Sigma}$, length n_T prediction vector $\mathbf{p}$.

1: Calculate class means for source data, source and target correspondence spectra: $\left\{\mathbf{S}_\mu, \mathbf{C}_\mu^S, \mathbf{C}_\mu^T\right\}$

2: Calculate the relation vectors for source class means and source and target correspondence spectra:
$\mathbf{S}_R \leftarrow [\mathrm{rel}(\mathbf{S}_\mu(j), \mathbf{S}_\mu)]_{j=1}^{k_S}$, $\mathbf{C}_R^S \leftarrow \left[\mathrm{rel}(\mathbf{C}_\mu^S(j), \mathbf{C}_\mu^S)\right]_{j=1}^{k_S}$,
$\mathbf{C}_R^T \leftarrow \left[\mathrm{rel}(\mathbf{C}_\mu^T(j), \mathbf{C}_\mu^T)\right]_{j=1}^{k_S}$

3: **for** $i = 1$ **to** n_T **do**

4: Calculate relation vector for current target pixel:
$\mathbf{r}_i \leftarrow \mathrm{rel}(\mathbf{T}(i), \mathbf{C}_\mu^T)$

5: $\mathbf{\Sigma}(i) \leftarrow \mathrm{relsim}(\mathbf{r}_i, \mathbf{S}_R) \cdot \mathrm{relsim}(\mathbf{r}_i, \mathbf{C}_R^S) \cdot \mathrm{relsim}(\mathbf{r}_i, \mathbf{C}_R^T)$

6: $$\mathbf{p}(i) = \begin{cases} \operatorname*{argmax}_j \mathbf{\Sigma}(i,j) & \text{if } \mathbf{\Sigma}(i,j) > \tau \\ 0 & \text{otherwise} \end{cases}$$

7: **end for**

RelThresh Algorithm

Algorithm 1 RelThresh

Input: $n_C \times k_S$ source and target correspondence likelihood matrices $\{\mathbf{L}^{cS}, \mathbf{L}^{cT}\}$, length n_C label vector $\mathbf{y}^C$, $y_i^C \in [1, k_S]$. τ search range $\tau_{\text{range}} = \left[\min(\mathbf{L}^T), \max(\mathbf{L}^T)\right]$. Total τ steps n_{step}

Output: RelTrans threshold τ_{best}.

1: Set best prediction $p_{\text{best}} = -\infty$, current threshold $\tau_{\text{cur}} = \max(\tau_{\text{range}})$, step size $\tau_{\text{step}} = \frac{\max(\tau_{\text{range}}) - \min(\tau_{\text{range}})}{n_{\text{step}}}$

2: **while** $\tau_{\text{cur}} > \min(\tau_{\text{range}})$ **do**

3:     Initialize $p \leftarrow 0$

4:     **for** $i = 1$ **to** n_C **do**

5:         **if** $\max_j \mathbf{L}^{cS}(i,j) > \tau_{\text{cur}}$ and $\max_j \mathbf{L}^{cT}(i,j) > \tau_{\text{cur}}$ **then**

            if $\operatorname{argmax}_j \mathbf{L}^{cT}(i,j) = y_i^C$ **then** $p \leftarrow p + 1$

6:     **if** $p > p_{\text{best}}$ **then** $p_{\text{best}} \leftarrow p$, $\tau_{\text{best}} \leftarrow \tau_{\text{cur}}$

7:     $\tau_{\text{cur}} = \tau_{\text{cur}} - \tau_{\text{step}}$

Evaluation Methodology

1. Choose 1000 labeled (reflectance) pixels via stratified sampling from source and target images
2. Define (up to) 50 control points for each source class
3. Scale all pixels by their L^2 norm to compensate for linear illumination effects [Pouch et al. 1990]
4. Split samples (50/50) into train / test sets (averaged over 5 folds)
   A. Calculate baseline transfer accuracy by classifying target samples using a minimum distance to class means classifier (MinDist)
   B. Classify target samples after applying RelTrans transform
   C. Flag n_{flag} target pixels via RelThresh procedure
   D. Flag n_{flag} worst MinDist predictions
5. Report transfer accuracy for MinDist and RelTrans before and after flagging