18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Manifold and Transfer Subspace Learning for Cross-Domain Vehicle Recognition in Dynamic Systems

Olga Mendoza-Schrock,
Senior Member *IEEE*
Sensors Directorate
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH, USA

Mateen M. Rizki
Dept. of Engineering and
Computer Science
Wright State University
Dayton, OH, USA

Vincent J. Velten,
Senior Member *IEEE*
Sensors Directorate
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH, USA

Abstract*—Transfer Subspace Learning has gained recent popularity in the literature for its ability to perform cross-dataset and cross-domain object recognition—enablers for data fusion. The ability to leverage existing data without the need for additional data collections is attractive for Automatic Target Recognition applications. For Automatic Target Recognition (or object assessment) applications, Transfer Subspace Learning is a game changer for dynamic systems, as it enables the incorporation of sparse and dynamically collected data into existing systems that utilize large, dense databases. A baseline Transfer Subspace Learning technique is the Transfer Fisher's Linear Discriminative Analysis, an approach based on Bregman divergence-based regularization. This paper modifies the implementation of the Transfer Fisher's Linear Discriminative Analysis technique by combining it with Manifold Learning and adjusting it to allow for a more systematic search of tuning parameters. Specifically, the Diffusion Map approach is utilized, a Manifold Learning approach based on heat diffusion. The modified technique is then utilized for cross-data and cross-domain electro-optical vehicle recognition.

Keywords—transfer learning; transfer subspace learning; electro-optical imaging; vehicle recognition, manifold learning

I. Introduction

Transfer Subspace Learning (TSL) has found a diverse range of applications, including cross-domain face recognition and cross-domain text categorization [1] [2]. TSL is an enabler for data fusion and an enabler for dynamic model building--an important component for Dynamic Data Driven Application Systems (DDDAS) [3]. The focus of this study is to extend the TSL framework by combining it with Manifold Learning for a robust Aided Target Recognition (AiTR) system capable of achieving high vehicle recognition rates. We seek to build an AiTR system that is robust to different operating conditions [4] including sensor modality, lighting conditions, shadows, weather, sensor type, terrain, image quality, and quality of metadata [5]. A robust AiTR system would leverage all '*similar*' data to recognize a new target of interest instead of having to collect large amounts of data on a new target before a recognition model could even be built. Data collections are resource intensive and can cost on the order of tens of thousands of dollars depending on the scope of the collection, the number of sensors utilized, and the complexity of the scenarios. Often, AiTR systems are built utilizing data from a particular data collection but suffer from dramatic performance loss when utilizing data from a different data collection or under real-world scenarios. By utilizing TSL, AiTRs can be extended to dynamic systems where they are more robust and applicable to scenarios outside of the ones from which they were developed.

We propose a combination of TSL and Manifold Learning to build the baseline dynamic AiTR. Furthermore, we modify the implementation of the Transfer Fisher's Linear Discriminative Analysis (TrFLDA) [1] technique to allow for a more systematic search of tuning parameters. We show results for cross-dataset vehicle recognition using electro optical (EO) imagery; specifically, we recognize a Toyota Avalon and a Nissan Sentra under vastly different lighting conditions. We also show results for cross-domain vehicle recognition utilizing the recognition of a Toyota Avalon and Nissan Sentra to recognize a Toyota Avalon and Mitsubishi Lancer.

This paper is organized as follows. Section II discusses Transfer Subspace Learning and Section III provides an overview of Manifold Learning. Section IV discusses the data utilized in the study and the experimental setup. Section V details experimental results. Section VI provides a discussion and our conclusions.

II. Transfer Subspace Learning

The purpose of Transfer Learning is to utilize information for recognition in one domain to recognize objects in a different but related domain. The seminal approaches include Transfer Adaboost [6] (TrAdaboost) which is an extension of the Adaboost algorithm for cross-dataset and cross-domain applications. It is common in the literature to report results using data from one data collection and reserve a part for training and a part for testing. However, cross-dataset recognition utilizes data from one data collection for training and data from a completely different data collection, under different operating conditions [4], for testing. The objects and their labels remain the same. For cross-domain recognition the objects and the labels are different but "similar", where the definition of similar is application dependent. In [1] similar meant that the objects were all faces. The datasets utilized did not share the same subjects but they were all faces under a set of restricted poses. In this study, for cross-dataset recognition we define *similar* to be the same set of sedans under two different lighting conditions and in cross-domain recognition we define *similar* to be any sedan.

Traditional transfer learning techniques, such as TrAdaboost, were next extended to Subspace Learning which resulted in a family of Transfer Subspace Learning (TSL) techniques [1] [7]. These techniques extend traditional subspace learning to account for the changes in distribution in the source and target domains. That is, in situations where independence is violated, TSL can be used under cross-dataset and cross-domain applications when the source and target data are not independent and identically distributed (i.i.d). Hence TSL techniques attempt to correct for this problem by either aligning the dataset or at least minimizing the difference in their distributions. This study focuses on TSL based on the Bregman Divergence-Based Regularization [1].

The assumptions of TSL based on the Bregman Divergence-Based Regularization are that the target domain has *one* labeled instance per class and that the amount of data in the target domain is less than the amount of data in the source domain. The goal is to find a subspace, W^*, that separates classes and aligns the distribution of the source and target data. This optimal subspace is then used to train labeled examples from the target domain by using the K-nearest neighbors (KNN) classifier [8]. We follow the technique of [1] and randomly select a sample from each of the classes in the target domain to classify the target data. In computing the optimal subspace, W^*, the key consideration is the trade-off between what can be learned from the current environment versus what is applicable to another environment. The variables that are important for separating classes in the source domain may differ from the variables that are important for separating classes in the target domain.

In this study, the Transfer Fisher's Linear Discriminative Analysis (TrFLDA) is utilized—one of six TSL approaches introduced in [1]. An assumption of TrFLDA, inherited from Fisher's Linear Discriminative Analysis (FLDA), is that the number of samples must be greater than or equal to the number of classes, plus the number of dimensions of the data [9]. A thorough explanation of the algorithm for TrFLDA is given in [1]. In this paper we provide a summary of the implementation in Algorithm I. In the original implementation [1] and associated example code [10], the optimization problem in Step One is expressed as

$$W^* = \arg\min_{W \in R^{Dxd}} F(W) + \lambda D_W\ (P_S \,\|\, P_T) \tag{1}$$

where $F(W)$ is the objective function for the particular Transfer Subspace Learning (TSL) approach, $D_W(P_S \parallel P_T)$ is the regularization term, and λ is the *absolute* weight assigned to the regularization term. Further, P_S and P_T are the probability density functions (PDFs) of the source data and the target data in the projected subspace, respectively.

For the TrFLDA analysis presented in this paper, F(W) is the objective function for FLDA given by

$$F(W) = \frac{Tr(W^T S_A W)}{Tr(W^T S_B W)} \tag{2}$$

where S_A is the standard *within class* distance and S_B is the standard *between class* distance for FLDA. The regularization term, which is the key to these techniques, is the Bregman divergence

$$D_W(P_S \parallel P_T) = \int [P_S(y) - P_T(y)]^2 dy \tag{3}$$

which is a measure of the difference between the distribution in the projected subspace, *W,* of the source data and the target data. The densities in the projected subspace are estimated using the kernel density estimation (KDE) technique [11]. There are several Bregman divergence measures to choose from such as mutual information, Kullback–Leibler (K-L) divergence, and Squared Euclidean distance (SED). For this effort, the SED was chosen as the baseline since it offers low computational cost [1] and in TrFLDA it may be more applicable than others since it appears it can better handle sparse data. In this study, the objective function of (1) is modified and is expressed as

$$W^* = \arg\min_{W \in R^{Dxd}} (1 - \lambda) F(W) + \lambda D_W\ (P_S \| P_T) \tag{4}$$

where $\lambda \in [0,1]$ is the *relative* weight assigned to the regularization term. This formulation of the objective function allows for a more systematic search of the optimal weight to assign to the regularization term. Further, note that whereas in the original implementation λ could range over the positive reals, the new implementation restricts λ to the interval [0,1]. This provides for a more systematic treatment as λ = 0 corresponds to the non-regularized approach, while λ = 1 corresponds to pure data alignment, with no learning from the original subspace being transferred. In the original implementation the latter scenario could be realized only asymptotically as $\lambda \to \infty$.

Algorithm I: Transfer Fisher's Linear Discriminative Analysis (TrFLDA)

Input: High-dimensional data in R^D (source and target data)
Output: $\boldsymbol{W}^* \in \boldsymbol{R}^{\boldsymbol{D} \times \boldsymbol{d}}$ (a linear mapping from R^D to R^d)
Output: Low-dimensional data in R^d (where $d \ll D$)
Output: Recognition rate using KNN

1. Find W^* by solving the problem
$$W^* = \arg\min_{W \in R^{D \times d}} (1 - \lambda) F(W) + \lambda D_W(P_S \parallel P_T)$$
using the gradient descent technique, as follows

1.1. Compute the initial linear mapping W_0 using the non-regularization subspace learning approach
$$W^* = \arg\min_{W \in R^{D \times d}} F(W)$$
For TrFLDA use Equation (2) for F(W)
1.2. Choose a value for $\lambda \in [0,1]$ (the regularization weight)
1.3. Choose a value for *K* (the maximum number of iterations on the gradient decent method)
1.4. Choose a value *h* (the optimization threshold)
1.5. Choose a value for $\eta(k)$ (the learning rate factor)
1.6. For each iteration $k \in \{1,2,\ldots K\}$, compute
$$W_{k+1} = W_k - \eta(k)\left((1-\lambda)\frac{\partial F(W)}{\partial W} + \lambda \frac{\partial D_W\ (P_S \| P_T)}{\partial W} \right)$$
For TrFLDA,

$$\frac{\partial F(W)}{\partial W} = \frac{2}{Tr(W^T S_B W)} S_A W - 2(Tr(W^T S_{\boldsymbol{B}} W))^2 Tr(W^T S_A W) S_B W$$

1.7. Terminate when the optimization threshold (h) or the maximum number of iterations (k) is reached
2. Construct the *reference set* from the target domain data by randomly selecting one sample per class
3. Apply the optimal subspace W^* to the reference set and the testing data from the target domain
4. Apply KNN using the reference set for training and the target domain data for testing
5. Calculate the recognition rate

III. MANIFOLD LEARNING

A. General manifold learning

Manifold learning involves finding the underlying structure of data to achieve non-linear dimension reduction. The goal of these techniques is to learn a mapping from the original high-dimensional data observation space to a lower-dimensional space that captures the underlying structure in the data. Manifold learning techniques are based on the assumption that the observed high dimensional data is parametrized by only a few degrees of freedom. These techniques evolved from Principal Component Analysis (PCA) [12], a linear method ill-suited for analyzing non-linear phenomena. Manifold learning techniques were created to overcome this limitation. Common techniques in manifold learning include Isomaps [13], Laplacian Eigenmaps [14], Local Linear Embedding [15], Multi-dimensional Scaling [16], and Diffusion Maps [17], to name a few. The focus of our study is on Diffusion Maps as explained below.

B. Diffusion maps

Diffusion maps are a non-linear dimension reduction technique introduced by Lafon *et al.* in [17], [18], [19], and [20]. Diffusion maps are of particular interest for ATR applications as the technique is robust to data fusion. That is, the input data can originate from sensors of different modalities. The technique is amenable to multi-sensor applications where data is collected from different sensors over the same area and of the same targets. The main benefits of the technique are that it is fast and robust to non-uniform sampling and noise. Two areas of active research are expanding the technique's ability to handle sparse sampling and reducing sensitivity to tuning parameters.

The diffusion maps technique derives a multi-scale, low-dimensional embedding from high-dimensional data by considering a random walk over a graph of the data. For a thorough explanation of the technique see [21]. However a simple explanation of the approach is given in Algorithm II. There are five decisions in the algorithm. The first choice is the similarity metric to be used to measure 'similarity' between two data points. This is, of course, application dependent and different similarity measure choices will lead to very different results. To simplify the implementation, the Euclidean distance is often employed as the measure of 'similarity' between data points but any symmetric, non-negative distance function can be utilized. The second choice is to select a kernel in order to construct the graph. For simplicity the Gaussian kernel is regularly employed, but again this specific choice is application dependent. The Gaussian kernel has proven to be a good choice having been used in a number of applications from gender classification [22] to vehicle classification [23]. The three remaining choices are all tuning parameters—σ, t, d. The parameter d is the number of dimensions in the low-dimensional space. The parameter σ is a positive scale parameter and controls the connectivity of the graph. This is the parameter that has been most studied and was the focus of the last publication by the authors in this area [21]. Lastly, the parameter t is the time in the random walk, i.e. the t^{th} step of the random walk—it is the exponent of the transition probability matrix.

Algorithm II: Diffusion Maps

Input: High-dimensional data in R^D, $X = \{x_1, x_2, ..., x_M\} \subset R^D, \sigma, t, M$

Output: Low-dimensional data in R^d, where d<<D, $Y = \{y_1, y_2, ..., y_M\} \subset R^d$

1. Normalize the high dimensional data
2. Compute the Euclidean distance between samples
3. Select a value for the positive scale parameter σ
4. Construct the weighted graph G matrix using the Gaussian kernel, $k_{ij} = \exp(-\| x_i - x_j \|^2 / 2\sigma^2)$
5. Compute the transition probability matrix $P_{ij} = k_{ij} / d_i$, where $d_i = \sum_{j=1}^{M} k_{ij}$
6. Select a value for t
7. Compute P^t
8. Apply singular value decomposition (SVD) to P^t to obtain its eigenvalues, $\{\lambda_l\}$, and eigenvectors $\{\psi_l\}$
9. Sort the eigenvalues and eigenvectors in descending order
10. Compute the diffusion map defined as $y_i = \Psi_t(x_i) = [\lambda_1^t \psi_1(i), ..., \lambda_d^t \psi_d(i)]^T$

To improve vehicle recognition performance, diffusion maps and transfer subspace learning (TSL) are combined in this study. According to [1] and the associated example implementation [10] principal component analysis is first used on the source and target domain data to reduce the dimensionality to a manageable number of dimensions. In that study the data utilized were raw images. In our study we utilize diffusion maps to reach a manageable number of dimensions.

IV. EXPERIMENTS

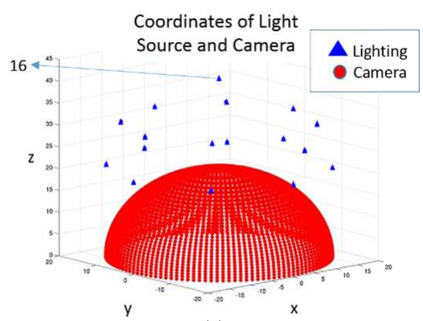
A. Electro-optical synthetic vehicle data domes

The data used in this study is a subset of the Electro-Optical Synthetic Civilian Vehicle Data Domes dataset published and maintained by the Air Force Research Laboratory (AFRL) Sensors Directorate and presented in [24].

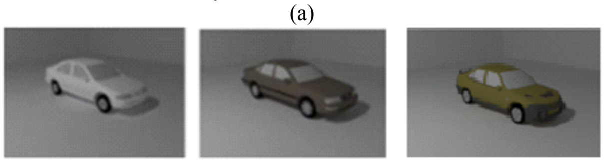
This dataset is unique as the vehicles in the dataset were all derived from three-dimensional (3D) point clouds of physically accurate vehicle models. Hence, although the data is synthetic, the models used for the vehicles were all derived from physically accurate dimensions. The complete dataset consists of ten vehicle types, but due to rendering errors which will be fixed in future iterations on the dataset, we only use four of the vehicles—Toyota Avalon, Jeep Cherokee, Honda Civic, and Nissan Sentra. The dataset is also unique in that images were generated using 17 different lighting conditions resulting in 3601 different poses with physically accurate shadows. The original data consists of 480 × 640 resolution color images. The poses were systematically captured every three degrees in azimuth and elevation—hence for a given elevation there are 120 poses. Figure 1 shows the distribution of camera and lighting location as well as a sample image for each of the vehicles. For a complete description of the data and directions on acquiring this data, see [24].

To reduce computational time, the images were down-sampled from 480 × 640 to a resolution of 160 × 213 by using a bi-cubic down-sampling scheme and were converted from color to grey-scale images.

For manifold learning we used the diffusion map technique described in Section III B. We kept the selection of the tuning parameters consistent with our previous work [21] where the kernel width parameter, σ, is set to seven and the diffusion time parameter, t, is set to one. For the number of dimensions, we selected M as 45 to be able to compare our results to the previous work which identified M=45 as having the highest classification rates.



(a)



Nissan Sentra Toyota Avalon Mitsubishi Lancer

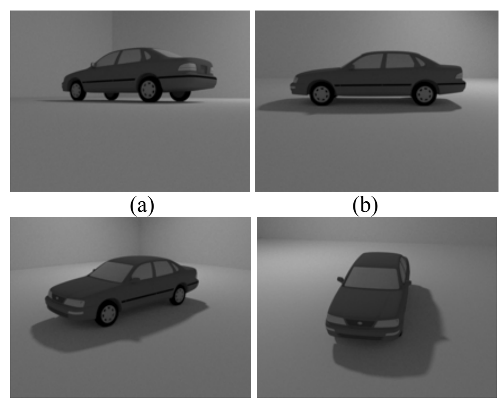
(b)

Fig. 1. (a) The distribution of lighting positions in blue triangles and camera positions in red circles; lighting condition 16 is highlighted as it is the nadir position at (0,0,41) (b) Sample vehicle images.

B. *Experimental design*

The datasets used in the experiments are from three different camera positions entitled Lighting Condition (LC) One, Nine, and 10. Data from LC one, at coordinates (14.142, 14.142, 21), is the source domain while data from LCs nine (12.247, 12.247, 31) and 10 (0, 17.321, 31) are target domains. The Toyota Avalon and the Nissan Sentra are utilized for the cross-dataset experiments while the Toyota Avalon, Nissan Sentra, and the Mitsubishi Lancer are utilized for the cross-domain experiments.

First a diffusion map was created for each of the different LCs utilizing the preprocessing and tuning parameters described in Section IV A. A swath of the data was utilized instead of the full data dome since recognition rates for the full data dome are near 50%. This rate is unacceptable as a baseline for a transfer learning study. Instead, a rather large data swath of 120 degrees in azimuth and 36 degrees in elevation is utilized for all experiments, specifically 141°–261° in azimuth (samples 48-88) and 0°–36° in elevation (samples 1-13). This swath size results in a dataset of 533 images. Example images from the swath are given in Fig. 2. By visual inspection, note the large variability in the poses and shadow in the swath. To quantify the difference between image swaths for the three different lighting conditions used in this study we utilize the average root mean square error (RMSE) [25] and give the results in TABLE I.



(a) (b) (c) (d)

Fig. 2. Example images from the Toyota Avalon swath used in the study, image (a) one, (b) 178, (c) 356, and (d) 533-the last image in the swath.

TABLE I. AVERAGE ROOT MEAN SQUARED ERROR (RMSE) IN PIXEL INTENSITY BETWEEN SOURCE AND TAGET DATASETS (533 IMAGES TOTAL PER DATASET)

Average RMSE	Source	Target
5.7211	Avalon LC-1	Avalon LC-9
5.7260	Sentra LC-1	Sentra LC-9
7.1563	Avalon LC-1	Avalon LC-10
7.3299	Sentra LC-1	Sentra LC-10
5.5932	Sentra LC-1	Lancer LC-1

The diffusion maps for LC one, nine, and 10 are given in Fig. 2. By visual inspection one can see that the diffusion maps for LC-1 and LC-10 are closer in shape than the diffusion maps for LC-1 and LC-9. The data from all the dimensions of the diffusion maps is the input to the Transfer Learning process. The diffusion maps are well-behaved in the sense that the first few dimension can be easily explained by analysis, the minor axes of the manifolds correspond to the 13 different elevation angles while the major axes of the manifolds correspond to the 40 different azimuth angles.

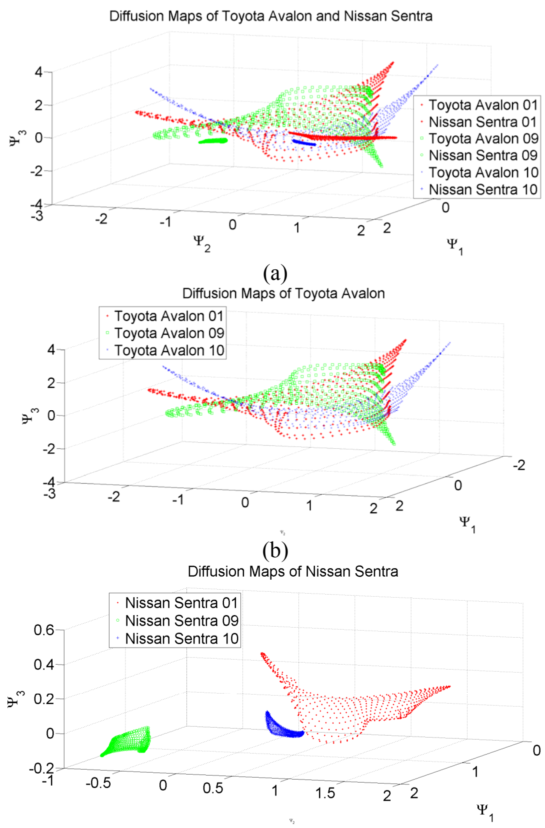


Fig. 3. The diffusion maps for the Toyota Avalon and Nissan Sentra in three different lighting conditions. (a) shows both targets (b) shows just the Toyota Avalon, and (c) displays just the Nissan Sentra.

V. Experimental Results

A. Cross-dataset vehicle recognition

The first experiment investigates the performance of Transfer Subspace Learning (TSL) via Transfer Fisher's Linear Discriminative Analysis (TrFLDA) for cross-dataset vehicle recognition. We seek high recognition rates of a Toyota Avalon and a Nissan Sentra under different lighting conditions. The source domain is Lighting Condition (LC) one and the target domain is LC-9. The TrFLDA assumption is satisfied, as the number of samples (533) is greater than the number of dimensions (45) plus the number of classes (2). A flow diagram for the cross-dataset experiments is depicted in Fig. 4.

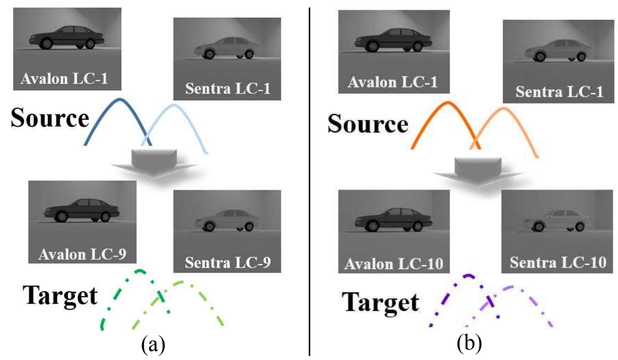


Fig. 4. Flow chart for cross-dataset experiments (a) TrFLDA for Avalon and Sentra from LC-1 to LC-9. (b) TrFLDA for Avalon and Sentra from LC-1 to LC-10. Solid lines indicate notional PDFs for source data and dashed lines indicate PDFs for target data

The baseline recognition rate for this experiment is 57.04% realized using FLDA on the source data, training the reference points and then using the reference points to classify the target data. In this particular experiment, since the labels are the same for both source and target domains, another baseline is calculated by utilizing the K-nearest neighbors classifier [8] and training with LC-1 and testing with LC-9. Using $k = 1$ the KNN results in a recognition rate of 75.7036% correctly classified instances, 167 missed Avalons, 92 missed Sentras, and a 0.5141 Kappa Statistic—a measure of how much better the classification is over random chance. TrFLDA results are compared to these baseline results. Recognition rates, the regularization term (λ), and the convergence iteration number for which the optimization converges for TrFLDA are given in TABLE II. For all three experiments described in Section IV, the number of maximum iterations (K) was set at 2000, the learning rate (η) was set at 0.05, and the threshold (h) was set at 0.00001. At each iteration, the recognition rate is calculated using the resulting projection and the KNN (k=1) classifier. An exhaustive search was utilized for λ values ranging from [0.00-1.00] at two decimal point increments. The best recognition rate of 74.2026% were found for values of λ varying from [0.2-0.3]. A selection of the results is shown in TABLE II.

TABLE II. Transfer Fisher's Linear Discriminative Analysis (TrFLDA) Results for Toyota Avalon and Nissan Sentra In Lighting Conditions 1 and 9

λ	Convergence Iteration	Recognition Rate
0	4	57.04%
0.1	112	74.02%
0.2	88	74.20%
0.3	78	74.20%
0.4	69	74.11%
0.5	62	74.11%

λ	Convergence Iteration	Recognition Rate
0.6	57	74.11%
0.7	52	74.11%
0.8	48	74.11%
0.9	45	74.11%
1	42	74.11%
0.21	87	74.20%
0.22	86	74.20%
0.23	85	74.20%
0.24	84	74.20%
0.25	83	74.20%
0.26	82	74.20%
0.27	80	74.20%
0.28	79	74.20%
0.29	79	74.20%

The second experiment is similar to the first but the source domain is LC-1 and the target domain is LC-10. The baseline for this experiment is a recognition rate of 93.996% realized using FLDA on the source data, training the reference points, then using the reference points to train the target data. Again, since the classes for the source and target domain are the same an additional baseline outside the transfer learning paradigm is calculated. That baseline performance for this experiment using KNN ($k = 1$) results in a recognition rate of 94.6529% correctly classified instances, 0.8931 kappa statistic, 57 missed Avalons, 0 missed Sentras. TrFLDA recognition rates, the regularization term (λ), and the convergence iteration number for which the optimization converges are given in TABLE III. An exhaustive search was utilized for λ values ranging from [0.000-1.000] at three decimal point increments. The best recognition rate of 96.3415% was found for the value of λ of 0.01. A selection of the results is shown in TABLE III.

TABLE III. TRANSFER FISHER'S LINEAR DISCRIMINATIVE ANALYSIS (TRFLDA) RESULTS FOR TOYOTA AVALON AND NISSAN SENTRA IN LIGHTING CONDITIONS 1 AND 10

λ	Convergence Iteration	Recognition Rate
0	4	94.00%
0.1	143	87.34%
0.2	88	86.96%
0.3	66	86.87%
0.4	53	87.15%
0.5	45	87.34%
0.6	40	87.43%
0.7	36	87.62%
0.8	32	87.80%
0.9	32	88.27%
1	24	89.59%
0.01	147	96.34%
0.02	258	94.37%
0.03	266	90.34%
0.04	241	88.84%
0.05	216	88.18%
0.06	196	87.71%
0.07	179	87.71%
0.08	165	87.62%
0.09	153	87.34%
0.005	136	95.87%
0.006	143	96.06%
0.007	145	96.15%
0.008	146	96.25%
0.009	147	96.25%
0.011	149	96.25%
0.012	152	96.06%
0.013	158	95.87%
0.014	168	95.78%
0.015	182	95.59%

B. *Cross-domain vehicle recognition*

The third experiment is a cross-domain experiment using diffusion maps and TrFLDA as in the cross-dataset experiments in Section IV.A. The source domain is the Toyota Avalon and the Nissan Sentra under lighting condition (LC) one and the target domain is the Toyota Avalon and Mitsubishi Lancer also under LC-1. The baseline for this experiment is a recognition rate of 0.5% realized using FLDA on the source data, training the reference points, then using the reference points to train the target data. Since the classes for the source and target domains differ, a baseline recognition rate outside the transfer learning paradigm cannot be calculated unlike the first two experiments. The regularization term, λ, the number of iterations it took to converge, and TrFLDA recognition rates are given in TABLE IV. An exhaustive search was utilized for λ values ranging from [0.0-1.0] at one decimal point increments. Given those results λ values of 0.85 and 0.95 were also explored. The best recognition rate of 65.1% was found for values of λ at 0.9 and 0.95. This recognition rate greatly outperforms the baseline by 15.1%. Future efforts will continue the exhaustive search and explore more efficient search methods to determine the optimal setting for λ.

TABLE IV. TRANSFER FISHER'S LINEAR DISCRIMINATIVE ANALYSIS (TRFLDA) RESULTS FOR SOURCE DATA OF TOYOTA AVALON AND NISSAN SENTRA AND TARGET DATA OF TOYOTA AVALON AND MITSUBSHI LANCER

λ	Convergence Iteration	Recognition Rate
0	4	50.00%
0.1	251	61.73%
0.2	306	64.53%
0.3	247	64.63%
0.4	205	64.54%
0.5	176	64.92%
0.6	155	64.82%
0.7	138	64.92%
0.8	125	65.01%
0.9	115	65.10%
1	106	65.01%
0.85	120	65.01%
0.95	110	65.10%

VI. DISCUSSION AND CONCLUSIONS

The results indicate that the Transfer Subspace Learning (TSL) techniques are sensitive to tuning parameters. The modification we propose to the implementation of Transfer Fisher's Linear Discriminative Analysis (TrFLDA) in Section II proves to be an improvement since the search space for λ is now bounded where $\lambda \in [0,1]$. Note that there is not a guarantee that the search spaces for λ and η are convex respectively so in certain applications an exhaustive search of the space would be necessary. Furthermore, a study into heuristics for these tuning parameters will be completed in future efforts.

For all three experiments the recognition rates using TrFLDA outperforms the baseline recognition rates using FLDA. Similarly, for the cross-dataset experiments the recognition rates using TrFLDA either outperform or match the recognition rates using KNN. In the case where TrFLDA doesn't outperform, in real-world applications TrFLDA would be preferred to KNN since the TrFLDA method only requires one labeled sample per class. Labeling data for use in recognition algorithms is expensive and manually intensive. As such, a technique with one labeled instance per class that is capable of matching the performance of a technique where all instances are labeled is a significant contribution towards a robust and sustainable Aided Target Recognition (AiTR) system.

The combination of TrFLDA with diffusion maps proved to be useful. One of the immediate benefits of the diffusion maps is first evident in the experimental design in Section IV.B. Notice how large the data swath is—the swath covers a span of 120 degrees in azimuth and 39 degrees in elevation. Fig. 2 displays four static images to help gain an appreciation for the variability in shadow and pose encompassed in the data swath. It is difficult to find a reference to other recognition systems that can handle such a diversity of target pose and still result in high recognition rates. As a comparison, the original TrFLDA study [1] utilized faces that varied in pose by 90 degrees in azimuth and zero degrees in elevation. Another benefit to the combination is the ability to explain the performance of TrFLDA based on the shape and scale of the diffusion maps. If the raw pixels were only considered, then based on TABLE I. a valid prediction would be that TrFLDA recognition rates would be higher for LC-1 and LC-9 than for LC-1 and LC-10. This is because LC-1 and LC-9 have a smaller difference in their RMSE measures and may be considered more '*similar*.' However, the opposite result is observed as shown in TABLE II. and TABLE III. – the TrFLDA recognition rates for LC-1 and LC-10 outperform the recognition rates for LC-1 and LC-9. This result is consistent with the visual inspection of the manifolds shown in Fig. 3. The diffusion map for LC-1 and LC-10 are actually closer in shape and proximity than the diffusion maps for LC-1 and LC-9. Hence, based on the diffusion maps, LC-1 and LC-10 are actually more '*similar*' than LC-1 and LC-9. Exploiting this observation and assessing its repeatability in other cross-dataset and cross-domain scenarios is the focus of the next stage of this research.

It is important to note that the current implementation of TrFLDA is dependent on the choice of reference point for each class which is randomly selected. Variations on the order of a percentage point as a function of choice of reference point have been observed in this study. A more thorough quantification of the effect of reference point selection is warranted. Furthermore, it is likely that reference point sets which include multiple samples per class would outperform the current methodology. This possibility will be explored in future work.

The dataset explored in this paper was for a single modality. The TSL framework can make important contributions to data fusion for data collections with multi-modal sensor types. One such example application would be training in the visual space and extending the learning to the infrared (IR) space using similar context. This will be explored in future work with datasets available in the fusion community literature (see e.g. [26]-[31]).

This paper presents the use of Transfer Subspace Learning (TSL) and Diffusion Maps for baseline vehicle recognition of fused datasets. Future work includes four areas. For the community the benchmark Electro-Optical Synthetic Vehicle Data Domes dataset will be extended and available for future analysis. Improvements include enclosing the vehicles in a circular containment instead of a square containment and component articulations of all vehicles including open hoods and open doors. Also, we seek to extend the TSL General Framework to include heuristics for the tuning parameters especial, λ, the regularization term, and, η, the learning rate factor. We will continue the study to include the other TSL approaches described in [1] including Transferred Principal Component Analysis (TrPCA), Transferred Locality Preserving Projections (TrLPP), and Transferred Discriminative Locality Alignment (TrDLA). Finally, we will extend the TSL framework to hierarchal transfer learning where for example the data learned on minivans and trucks can be extended to sedans given there is a hierarchal relationship between larger vehicles and smaller vehicles.

ACKNOWLEDGMENT

The authors would like to thank Mr. Edmund G. Zelnio, Dr. Erik Blasch, and Dr. Jared Culbertson from the US Air Force Research Laboratory (AFRL) and Dr. Michael L. Raymer, Dr. Frederick Garber, Dr. John C. Gallagher from Wright State University for their insight and guidance.

REFERENCES

[1] S. Si, T. Dacheng Tao, B. Gen, "Bregman Divergence-Based Regularization for Transfer Subspace Learning," *Knowledge and Data Engineering*, IEEE Transactions on , vol.22, no.7, pp.929-942, July 2010.

[2] S. Yang, M. Lin, C. Hou, C. Zhang, and Y. Wu,. "A general framework for transfer sparse subspace learning." *Neural Computing and Applications* 21.7, pp. 1801-1817, 2012.

[3] E. Blasch, G. Seetharaman, F. Darema '"Dynamic Data Driven Applications Systems (DDDAS) modeling for automatic target recognition ", Proc. SPIE 8744, Automatic Target Recognition XXIII, 87440J, May 20, 2013.

[4] B. Kahler, E. Blasch, and L. Goodwon, "Operating condition modeling for ATR fusion assessment", Proc. SPIE 6571, Multisensor, Multisource

Information Fusion: Architectures, Algorithms, and Applications 2007, April 9, 2007.

[5] M. Bryant, P. Johnson, B. M. Kent, M. Nowak, and S. Rogers, “Layered Sensing: Its Def., Attributes, and Guiding Principles for AFRL Strat. Tech. Dev.”, http://www.wpafb.af.mil/shared/media/document/AFD-080820-005.pdf, retrieved February 2015.

[6] W. Dai, Q. Yang, G. R. Xue, and Y. Yu. "Boosting for transfer learning." In *Proceedings of the 24th international conference on Machine learning*, pp. 193-200. ACM, 2007.

[7] S. Yang, M. Lin, C. Hou, C. Zhang, Y. Wu, "A general framework for transfer sparse subspace learning." Neural Computing and Applications 21.7, pp. 1801-1817, 2012.

[8] S. Lloyd. "Least squares quantization in PCM." *Information Theory*, IEEE Transactions on 28.2, pp. 129-137, 1982.

[9] A.M. Martinez and A.C. Kak, “PCA versus LDA,” IEEE Trans. Pattern Analysis and Machine Intelligence , vol. 23, no. 2, pp. 228– 233, Feb. 2001.

[10] S. Si, Bregman Divergence Based Regularization for Transfer Subspace Learning Example Code, http://www.cs.utexas.edu/~ssi/, retrieved January 2015.

[11] M.P. Wand and M.C. Jones. *Kernel Smoothing*. Chapman and Hall, 1995.

[12] M. Wall, A. Rechtsteiner, and L. Rocha, “Singular value decomposition and principal component analysis,” [A Practical Approach to Microarray Data Analysis], Springer, US, pp. 91-109, 2003.

[13] J. B. Tenenbaum, V. D. Silva, J. C. Langford., “A global geometric framework for nonlinear dimensionality reduction,” *Science*, 290, pp. 2319-2323, 2000.

[14] M. Belkin and P. Niyogi, “Laplacian eigenmaps for nonlinear dimensionality reduction and data representation.” *Neural Computation*, 15(6), pp. 1373-1396, 2003.

[15] S. Roweis and L. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” *Science*, 290, pp. 2323-2326, 2000.

[16] T. F. Cox and A. A. Cox., [Mulidimensional scaling], CRC Press, US, 2001.

[17] R. R. Coifman et al. , “Geometric Diffusion as a Tool for Harmonic Analysis and Structure Definition of Data: Diffusion Maps,” Proc. National Academy of Sciences, 102(21), pp. 7426-7341, 2005.

[18] S. Lafon, Diffusion Maps and Geometric Harmonics, Ph.D. Dissertation, Yale University, 2004.

[19] R. R. Coifman, S. Lafon, “Diffusion Maps,” *Applied and Computational Harmonic Analysis*, Vol. 21, 5-40, 2006.

[20] S. Lafon, Y. Keller, and R. R. Coifman, “Data Fusion and Multicue Data Matching by Diffusion Maps”, IEEE Trans. *Pattern Anal. Mach. Intell*., Vol. 28, no. 11, pp. 1784-1797, 2006.

[21] J. Ramirez, and O. Mendoza-Schrock, "Diffusion maps for exploring electro-optical synthetic vehicle image data," Aerospace and Electronics Conference (NAECON), 2012 IEEE National , vol., no., pp.126-133, 25-27 July 2012.

[22] O. Mendoza-Schrock, J. Patrick, G. Arnold, and M. Ferrara, ‘Applying manifold learning techniques to the CAESAR database.’ Evolutionary and Bio-Inspired Computation: Theory and Applications IV. Proceedings of the SPIE, Volume 7704, pp. 77040O-77040O-11, April 2010.

[23] Y. S. Bhat and G. Arnold, “Diffusion Maps and Radar Data Analysis”, Algorithms for Synthetic Aperture Radar Imagery XIV, Proc. of SPIE, Vol. 6568, April 2007.

[24] R. L. Price, J. Ramirez, T. V. Rovito, and O. Mendoza-Schrock, "Electro-optical synthetic civilian vehicle data domes," *Aerospace and Electronics Conference (NAECON), 2012 IEEE National* , vol., no., pp.140-143, 25-27 July 2012.

[25] T. Chai and R. R. Draxler. "Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature." Geoscientific Model Development 7.3, pp. 1247-1250, 2014.

[26] H. Ling, L. Bai, E. Blasch, and X. Mei, "Robust infrared vehicle tracking across target pose change using L 1 regularization." Information Fusion (FUSION), 2010 13th Conference on. IEEE, 2010.

[27] Y. Wu, E. Blasch, G. Chen, L. Bai, and H. Ling, "Multiple source data fusion via sparse representenation for robust visual tracking." Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on. IEEE, 2011.

[28] P. Liang, G. Teodoro, H. Ling, E. Blasch, G. Chen, and L. Bai. "Multiple kernel learning for vehicle detection in wide area motion imagery." In Information Fusion (FUSION), 2012 15th International Conference on, pp. 1629-1636. IEEE, 2012.

[29] E. Blasch, P. Maupin, and A-L. Jousselme. "Sensor-based allocation for path planning and area coverage using UGSs." In Aerospace and Electronics Conference (NAECON), Proceedings of the IEEE 2010 National, pp. 361-368. IEEE, 2010.

[30] C. Yang, I. Kadar, E. Blasch, and M. Bakich. "Comparison of information theoretic divergences for sensor management." In SPIE Defense, Security, and Sensing, pp. 80500C-80500C. International Society for Optics and Photonics, 2011.

[31] E. Blasch, C. Banas, M. Paul, B. Bussjager, and G. Seetharaman. "Pattern activity clustering and evaluation (PACE)." In SPIE Defense, Security, and Sensing, pp. 84020C-84020C. International Society for Optics and Photonics, 2012.