A Spatio-temporal Extension to Isomap Nonlinear Dimension Reduction

Odest Chadwicke Jenkins   CJENKINS@USC.EDU
Maja J Matarić   MATARIC@USC.EDU
Robotics Research Laboratory, Center for Robotics and Embedded Systems, Computer Science Department,
University of Southern California, 941 W. 37th Place, Los Angeles, CA 90089 USA

Abstract

We present an extension of Isomap nonlinear dimension reduction (Tenenbaum et al., 2000) for data with both spatial and temporal relationships. Our method, ST-Isomap, augments the existing Isomap framework to consider temporal relationships in local neighborhoods that can be propagated globally via a shortest-path mechanism. Two instantiations of ST-Isomap are presented for sequentially continuous and segmented data. Results from applying ST-Isomap to real-world data collected from human motion performance and humanoid robot teleoperation are also presented.

1. Introduction

The process of uncovering structure underlying unlabeled data is a challenging endeavor in unsupervised learning. Recently, several methods have been proposed to address this problem through dimension reduction from pairwise relationships. These include global techniques (e.g., Kernel PCA (Schölkopf et al., 1998), Isomap (Tenenbaum et al., 2000)), local techniques (e.g., Locally Linear Embedding (Roweis & Saul, 2000), Manifold Charting (Brand, 2002)), and spectral clustering (Ng et al., 2001). While these pairwise methods have exhibited great potential, several issues remain largely unaddressed, such as dealing with out-of-sample points (Bengio et al., 2003) and temporal dependencies within data.

Motivated by analyzing human and humanoid robot motion, we propose an a extension to Isomap for data with both spatial and temporal relationships. Two

versions of our *spatio-temporal Isomap (ST-Isomap)* are presented for continuous and segmented input data with sequential temporal ordering. Continuous ST-Isomap is suited for uncovering spatio-temporal manifolds of data exhibiting *temporal coherence*, where sequentially adjacent samples are incrementally different. Segmented ST-Isomap is suited for uncovering spatio-temporal clusters in segmented data, where the input data is prepartitioned. Our ST-Isomap method is validated with empirical results from applications to humanoid robot sensory data from teleoperation and multi-activity human motion capture data.

2. Spatio-temporal Dimension Reduction

The success of techniques mentioned in Section 1 is due largely to leveraging estimated spatial relationships between data pairs. Such methods are able to uncover global spatial relationships in data through local kernels, models, or neighborhoods about each point. For addressing temporal relationships, however, these techniques must be able to perform:

- *proximal disambiguation* of spatially proximal data points in the input space that are structurally different;

- *distal correspondence* of spatially distal data points in the input space that share common structure;

in order to uncover spatio-temporal structure. Our aim is to define dimension reduction techniques that perform proximal disambiguation and distal correspondence such that spatio-temporal structure becomes apparent.

Our notions of proximal disambiguation and distal correspondence are illustrated in Figure 1 with respect to three arm waving motions. In the left panel, the two
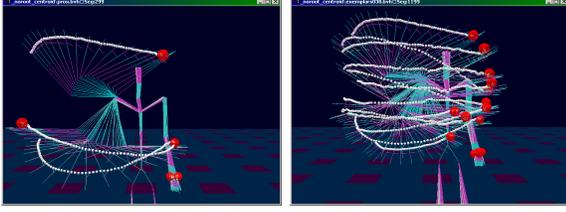
Figure 1. An illustration of proximal disambiguation and distal correspondence. (Left) Three waving motions with hand trajectories shown as a dotted trail. The beginning of each trajectory is marked with a large sphere. (Right) A set of exemplars connected the low and high waving motions through similarly structured motions.

low waving motions are relatively proximal in joint angle space but are structurally different due to moving in opposite directions. In constrast, the low and high motion waving the same direction are structurally corresponding but are distal in joint angle space.

We consider the case of waving back and forth at various heights as a single continuous performance. For such motion, we would expect to uncover a circular "loop" structure. Each iteration of the loop indicates on back and forth wave. For disambiguation, consider that proximal arm postures in joint angle space could be encountered during waves in both directions. Being from different underlying behaviors (e.g., "wave left" and "wave right") , such points should be separated to distal locations in the embedding space. In contrast, distal postures encountered during the high wave correspond to equivalent progress through a low wave. These corresponding motions should be placed into proximity in the resulting embedding, flattening the variations of waving to a single curve.

Because proximal disambiguation and distal correspondence are pair-based concepts, our approach is to augment existing pairwise dimension reduction for spatio-temporal relationships. Isomap is our primary focus for spatio-temporal augmentation, given its idependence on a specific measure of distance (due to its foundation in multidimensional scaling). In contrast, LLE is reliant upon weights and locally linear models that are inherently spatial and difficult to extend for other factors. Kernel PCA and kernel spectral clustering use local kernels that are applied globally from each data point, but may not be globally appropriate. Methods that abstract input data into intermediate models, such as in work by Teh and Roweis (Teh & Roweis, 2002) and Brand (Brand, 2002), could also be amenable to temporal extension at different levels of resolution. Temporal Kohonen Maps (Varsta et al., 2001) are suited to uncover structure in spatio-

temporal data, but require *a priori* specification of expected structure topology.

3. Spatio-temporal Isomap

The general framework for dimension reduction using Isomap is a batch three-step procedure for embedding a full matrix of geodesic distances. The distance matrix is computed by propagating local distances globally through an all-pairs shortest paths algorithm. Our extension, termed *spatio-temporal Isomap (ST-Isomap)*, retains this framework, inserting additional steps for temporal windowing and temporal augmentation. Assuming input data as samples from a continuous process, the general procedure for ST-Isomap is specified as follows, with extensions to Isomap indicated in bold:

1. **windowing of the input data into temporal blocks** S;

2. compute sparse distance matrix D^l from local neighborhoods $\texttt{nbhd}(S_i)$ about each point S_i using Euclidean distance;

3. **locally identify** *common temporal neighbors* $\texttt{CTN}(S_i)$ **of each point** S_i **as either** *local segmented common temporal neighbors* $\texttt{LSCTN}(S_i)$ **or** *K-nearest nontrivial neighbors* $\textit{KNTN}()$:

4. **reduce distances in** $D^l_{S_i,S_j}$ **between points with common and adjacent temporal relationships**:

$$D^0_{S_i,S_j} = \qquad\qquad (1)$$

$$\begin{cases} D^l_{S_i,S_j}/(c_{\text{CTN}}c_{\text{ATN}}) & \text{if } S_j \in \texttt{CTN}(S_i) \\ & \text{and } j = i+1 \\ D^l_{S_i,S_j}/c_{\text{CTN}} & \text{if } S_j \in \texttt{CTN}(S_i) \\ D^l_{S_i,S_j}/c_{\text{ATN}} & \text{if } j = i+1 \\ \texttt{penalty}(S_i,S_j) & \text{otherwise} \end{cases}$$

5. complete D^0 into full all-pairs shortest-path distance matrix $D = D^g$ (Dijkstra's algorithm), such that $g \geq |S|$:

$$D^g_{i,j} = \begin{cases} D^0_{i,j} & g = 0 \\ \texttt{min}(D^{g-1}_{i,j}, D^{g-1}_{i,k} + D^{g-1}_{k,j}) & g \geq 1 \end{cases} \quad (2)$$

6. embed D into d_e-dimensional embedding space through MDS such that:

$$E = |D^g - D_e|_{L^2} \qquad (3)$$

where `nbhd()` are the local neighbors of given segment, c_{CTN} and c_{ATN} are constants for increasing similarity between common and adjacent temporal neighbors, D_e is the matrix of Euclidean distances in the embedding, and $\|A\|$ is the L^2 matrix norm of A. `penalty`(S_i, S_j) is a function that determines the distance between a pair with no temporal relationship, typically set as $D^l_{S_i,S_j}$.

The first step, temporal windowing, serves to provide a temporal history for each data point. The result from windowing is a sequentially ordered set of data points S. This windowing is an initial (but not complete) means for temporal disambiguation. If we consider temporal windows at each point, we assume S maintains temporal coherence of the underlying process between sequentially adjacent points. This "continuous" data is suited for continuous ST-Isomap. If temporal windows are non-overlapping, temporal coherence is not assumed and segmented ST-Isomap is appropriate.

The third step in the procedure serves to establish *hard* spatio-temporal correspondences between proximal data pairs. Given $S_j \in$ `nbhd`(S_i), S_j is in the set of *common temporal neighbors (CTN)* that are local to S_i if a spatio-temporal correspondence between the pair is determined. `CTN()` can be defined by a variety of metrics. Described later in this section, we have chosen `KNTN()` for continuous data and `LSCTN()` for segmented data. CTN identified in this step are *local* individual neighborhoods.

In the fourth step, distances between data pairs with spatio-temporal relationships are reduced to accentuate their similarity. We consider two types of temporal relationships between a data pair, CTN and *adjacent temporal neighbors (ATN)*. These relationships allow for the construction a matrix of spatio-temporal similarities D^0 that will be globally propagated through all-pairs shortest-path computation. ATN are adjacent points in the sequential order of S. Elements in D^0 for distances between ATN explicitly establishes the temporal order in the data. Additionally, these elements ensure a single connected component will result in the all-pairs shortest-matrix D^g. Distances are greatly reduced between data pairs with CTN relationships by a constant factor of c_{CTN}. As the value of c_{CTN} increases, the distance between data pairs with spatio-temporal correspondences decreases and their similarity increases.

We consider CTN to be symmetric and transitive:

$$S_j \in \text{CTN}(S_i) \Leftrightarrow S_j \in \text{CTN}(S_i) \qquad (4)$$

$$S_j \in \text{CTN}_{\texttt{global}}(S_i) \Leftarrow \qquad (5)$$
$$S_j \in \text{CTN}_{\texttt{local}}(S_k) \text{ and } S_k \in \text{CTN}_{\texttt{local}}(S_i)$$

CTN transitivity (illustrated in Figure 3) is enforced through the shortest-path mechanism and not explicitly represented. Given a significantly large value c_{CTN}, distances between CTN-pairs will be reduced such that all pairs connected by a CTN-path can be identified as a *CTN component*. CTN components are identified assuming that local neighbors with CTN relationships will have significantly smaller distances in D^0 than non-CTN neighbors. In D^g, consequently, a point S_i will be more proximal to $S_j \in \text{CTN}(S_i)$ in its CTN component than any inter-component point $S_k \notin \text{CTN}(S_i)$:

$$S_j \in \text{CTN}(S_i) \text{ and } S_k \notin \text{CTN}(S_i) \Rightarrow \qquad (6)$$
$$D^g_{S_i,S_j} < D^g_{S_i,S_k}$$

By reducing local CTN distances by c_{CTN}, any two points S_i and S_j with a connecting path of CTN correspondences should have a shortest path of all reduced-distance edges. Any point S_k whose shortest path incurs an edge outside of the CTN component will put its distance to S_i outside of proximity. Beyond the scope of this paper, soft correspondences could be incorporated into ST-Isomap through using c_{CTN} as a variable weighing the degree of spatio-temporal similarity.

Once the full spatio-temporal distance matrix D^g has been generated, MDS is performed to produce the d_e-dimensional embedding. From this point, the same process as Isomap is performed. This process uses MDS to realize coordinates such that the distances of D^g are preserved. For this paper, we treat MDS as a "black box" that could be performed using a variety of techniques (Cox & Cox, 1994). Additionally, he embedding dimensionality can be selected by identifying the "elbow" of residual variance, as with Isomap.

We loosely term the structure produced by ST-Isomap a spatio-temporal manifold, an example of which is shown in Figure 2. This manifold is a structurally a 1-manifold curve in the embedding space. Each location on this curve is representative of a certain point of temporal progress along the spatio-temporal process. Each location on the curve also encapsulates all of the spatial variations representing a certain fixed spatio-temporal progress. Thus, a diverse set of spatial variations corresponding to the same spatio-temporal progress is collapsed into a single location in the embedded manifold.
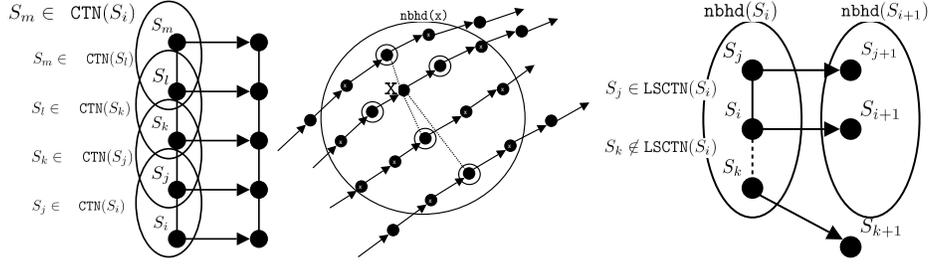
Figure 3. Illustrations of CTN transitivity corresponding S_i and S_m (left), K-nearest non-trivial neighbors of x (center), and local segmented common temporal neighbors of S_i (right). The KNTN of x are circled and trivial matches are marked with an "x". of distal correspondence of two points S_i and S_m through CTN transitivity.
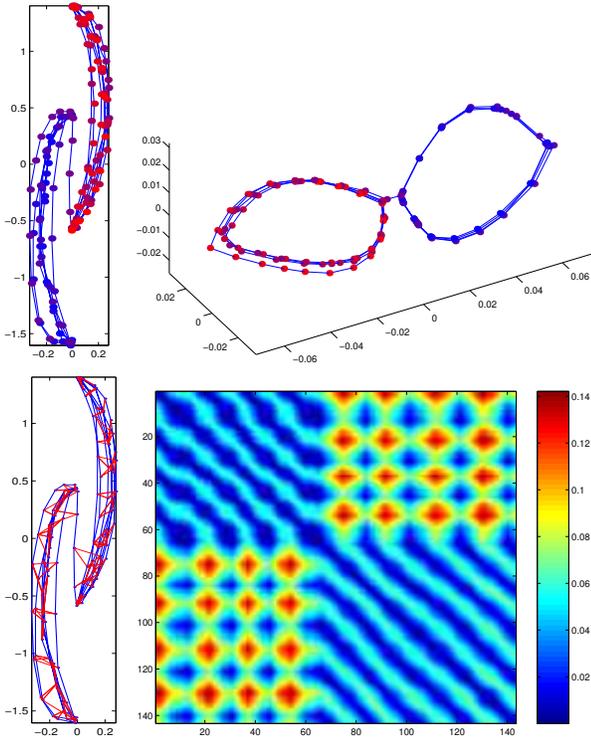

Figure 2. A temporal variation on the "two moons" example (Zhou et al., 2003) illustrating a spatio-temporal manifold. (Top left) 2D input data (drawn from blue to red with respect to time) collected from mouse movements moving up and down one moon and transitioning to move up and down the second moon. (Top right) The resulting continuous ST-Isomap embedding (KNTN = 3), producing two loops connected by a transition. (Bottom left) Hard spatio-temporal correspondences (shown by red lines) established by ST-Isomap in the input data. (Bottom right) Distance matrix produced by ST-Isomap.

3.1. Continuous ST-Isomap

We now describe the use of ST-Isomap for continuous data (i.e., data exhibiting temporal coherence). Con-

tinuous ST-Isomap assumes temporal disambiguation occurs by windowing over some horizon ϵ_d from each data point. Windowing in this manner is a means to include the velocity of the underlying process in the similarity matrix. CTN correspondences with respect to each data point are determined as its *K-nearest non-trivial neighbors (KNTN)* using Euclidean distance. Our notion of KNTN was inspired by Chiu et al. (Chiu et al., 2003), who define the concept of *trivial matches* in data mining for univariate time-series. Diverging slightly from their definition, we consider a point S_j to be a *nontrivial match* within the local neighborhood of a point S_i if it is closest matching point on its trajectory through the neighborhood (Figure 3):

$$S_j \in \texttt{KNTN}(S_i) \Leftrightarrow \qquad (7)$$
$$j = i + 1 \text{ or}$$
$$i \neq j \text{ and } D_{i,j}^l \leq D_{i,k}^l, k - \epsilon_w \leq k \leq k + \epsilon_w$$

where $(2\epsilon_w) + 1$ is the length of a trivial match window centered on point S_k. The KNTN of S_i are its K nearest nontrivial matches based on Euclidean distance. Given K, a data point $S_j \in KNTN(S_i, K) \Rightarrow S_i \in CTN(S_j)$ for continuous ST-Isomap.

The thought driving KNTN is that a large number of neighbors are expected be spatio-temporally similar to a point S_i. However, the bulk of these neighbors are redundant correspondences generated from a smaller number of trajectories passing through the neighborhood. KNTN effectively aims to find the best matching neighbors from each individual trajectory.

3.2. Segmented ST-Isomap

As with Isomap, a significant problem in using continuous ST-Isomap is its computational sensitivity to the number of samples N in the input data, requiring the storage, shortest-path computation, and eigende-

composition of an $N \times N$ matrix. Because input data are related in time by an underlying spatio-temporal process, the input can be partitioned into N_s non-overlapping segments, where $N_s \ll N$.

By forming S as N_s segments, ST-Isomap can be applied to larger input datasets. Because the Isomap framework is relatively insensitive to high-dimensional data, ST-Isomap is better equipped to handle a smaller number of N_s segments of higher dimension $d \times l$ as input rather than a larger number N of samples with lower dimension d. As discussed in (Jenkins, 2003), abstracting input samples into segments assumes mechanisms for segmentation and time normalization. In addition, segmented data requires a different definition for CTN correspondences. This new definition is necessary because sequentially adjacent points use to establish CTN-pairs may be distal in the input space.

Towards this end, we define *local segmented common temporal neighbors (LSCTN)* as (Figure 3):

$$S_j \in \texttt{LSCTN}(S_i) \Leftrightarrow \qquad (8)$$
$$S_j \in \texttt{nbhd}(S_i) \text{ and } S_{j+1} \in \texttt{nbhd}(S_{j+1})$$

The intuition driving SCTN is that a pair of points are spatio-temporally similar if they are spatially similar and the points they transition to are also spatially similar. More specifically, two segments S_i and S_j sharing a common spatio-temporal structure A will always be followed by segments S_{i+1} and S_{j+1} also sharing a common structure B, forming a temporal structure $A \rightarrow B$. Given a sufficiently large $c_{\texttt{CTN}}$, the resulting embedding will place points of an SCTN component into *clusterable proximity*, yielding separable clusters. We recommend the use of "sweep-and-prune" clustering (Cohen et al., 1995) into axis-aligned bounding boxes in such embeddings. Sweep-and-prune clustering uses a threshold distance on data projections to each axis for partitioning, avoiding the estimation of K cluster cardinality (Jain & Dubes, 1988).

We also note that segmentation presents a particularly challenging "chicken-and-egg" problem as there is no definitive general ground-truth domain-independent models or mechanisms to guide the abstraction of the input samples. In order to produce a structurally appropriate embedding, the segments produced from the input samples must be *consistent* (i.e., similar input intervals produce similar segments) and *atomic* (i.e., the user considers each segment to contain a conceptually and/or meaningfully indivisible performance/subsequence of the input data).

3.3. Connections to Hidden Markov Models

The result from clustering is a temporal process structure similar to Hidden Markov Models (Rabiner, 1989). A spatio-temporal process is uncovered as a structure in the form of "$\ldots \rightarrow A \rightarrow B \rightarrow C \rightarrow \ldots$", where A, B, and C are clusters in the embedding. Each cluster can be thought of as a latent variable grouping observed spatial variations on a spatio-temporal structure. The initial state probabilities of a state A can be computed via the normalized population of its cluster. Additionally, members of a cluster for A are found implicitly using the members of B, indicative of the transitional relationship between A and B. Thus, transition probability from states A to B can be the transition count from one cluster A to cluster B normalized by the number of transitions from A.

4. Results

In this section, we present results from applying MATLAB implementations, based on code provided by the authors of Isomap [1], of continuous and segmented ST-Isomap to human and robot data acquired from real-world performances.

4.1. Embedding Robonaut Sensor Data

To evaluate its functionality, we applied continuous ST-Isomap to sensory data produced from teleoperation of the NASA Robonaut (Ambrose et al., 2000) (Figure 4). Robonaut is a humanoid robot with upper body actuation of arms with 7 and 12 degrees of freedom (DOF) in each arm and hand, respectively. Additionally, various tactile and force sensors placed throughout Robonaut's upper body and hands.

For our input data[2], Robonaut was teleoperated to perform 5 trials for grasping of a horizontally oriented wrench from a starting rest posture. The wrench was placed at various locations in Robonaut's reachable space. During each trial, Robonaut published its sensor and motor actuation data at approximately 10Hz as a 110-dimensional vector. Motor actuation values were zeroed-out of this data, producing 460 frames of 57-dimensional vectors for each trial. Sensor columns were mean subtracted and normalized into a fixed range. Frames across all trials were concate-

---

[1]Thank you!

[2]The Robonaut teleoperation data were graciously provided by Alan Peters of Vanderbilt University and the Robonaut team at the Johnson Space Center. This data is temporarily available at http://robotics.usc.edu/ cjenkins/sensordata.zip and will persist at (Howard & Roy, 2003)
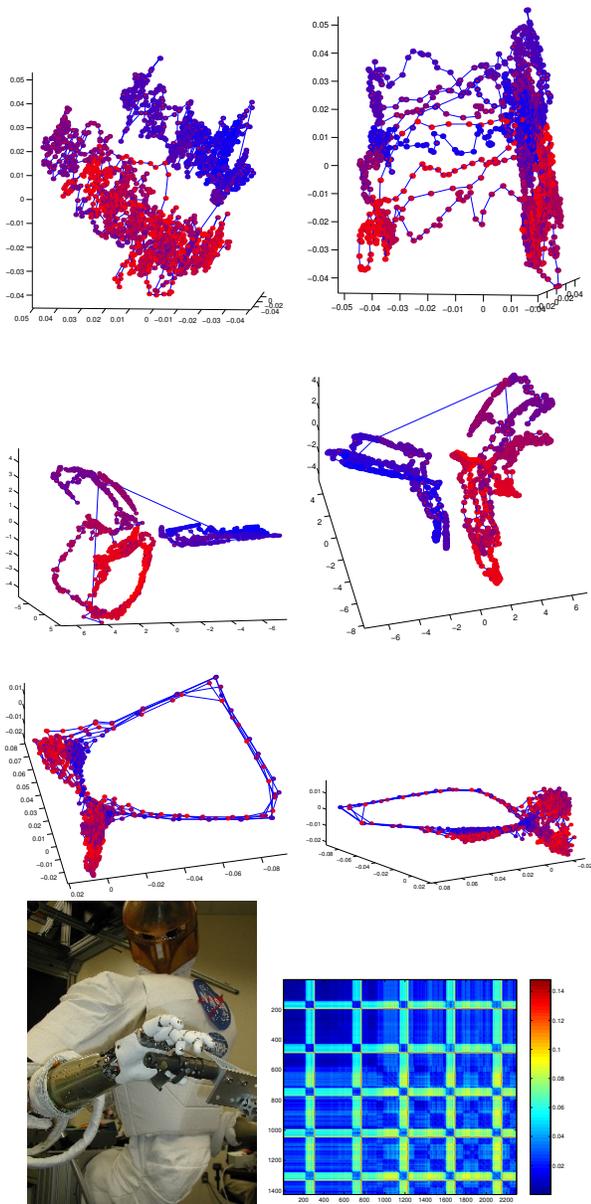
repeatedly performing the same behavior can be observed. However, both the PCA and Isomap embeddings are unable to capture the spatio-temporal structure of this data, resulting in embeddings that require significant deciphering. In contrast, ST-Isomap is able to capture the spatio-temporal process underlying this data as a curve with two large clusters. The looping structure contains 5 loops, indicative of the 5 grasp trials, with the two clusters representative of the idle time spent in the resting and grasping positions.

4.2. Embedding Multi-activity Human Motion

To evaluate its functionality, two time-series of kinematic motion were given as input into segmented ST-Isomap. This kinematic motion data[3] were acquired of a human subject performing multiple scripted high-level activities, including various dancing, punching, and arm waving behaviors (Input Motion 1) and two arm reaching to various locations (Input Motion 2). This data contain 22,549 and 9,394 frames, for Input Motions 1 and 2 respectively, of 42 kinematic DOF for rotations about joints of the arms and legs. Without segmentation, processing this motion data would be intractable for our MATLAB routine (capable of handling approximately 2500 points). This data was partitioned into 226 and 64 segments using Kinematic Centroid Segmentation (Jenkins, 2003), an automated procedure that treats limbs as pendulums and looks for limb "swings".

Embeddings and clusterings produced by segmented ST-Isomap from the motion segment data are shown in Figure 5. The reaching motion (Input Motion 2), the easiest of the motions to visually interpret, was mostly segmented into two primitive-level behaviors, "reach to position" and "return to rest posture". The PCA embedding illustrates the distinction of the two alternating behaviors, but is unable to distally correspond points of the same underlying behavior. The embedding produced by segmented ST-Isomap, however, is able to collapse the points into two distinct clusters with dominate populations. Spurious data points are also present in the ST-Isomap embedding that are a result of segments of idle motion due to resting between reaches. Because such resting periods are underrepresented in the input data, distal correspondence was unable to build a CTN component for this behavior.

Figure 4. 3D embeddings of 57D Robonaut sensor data from PCA (top row), Isomap using 20 K-NN (top middle row), and continuous ST-Isomap using 3 KNTN (bottom middle row). An image of the NASA Robonaut and the distance matrix from continuous ST-Isomap (bottom). In each plot, sequentially adjacent points are connected by a blue line and temporal order is color coded from blue to red.

nated together to form an input data set of 2300 57-dimensional samples.

Results of applying continuous ST-Isomap to Robonaut sensor data are shown in Figure 4. In the PCA embedding, the looping nature of the teleoperation

---

[3]The motion capture data were obtained with a Vicon optical motion capture system and graciously provided by Jessica Hodgins at Carnegie Mellon University. These motions are available for viewing at *http://robotics.usc.edu/ cjenkins/motionmodules/.*
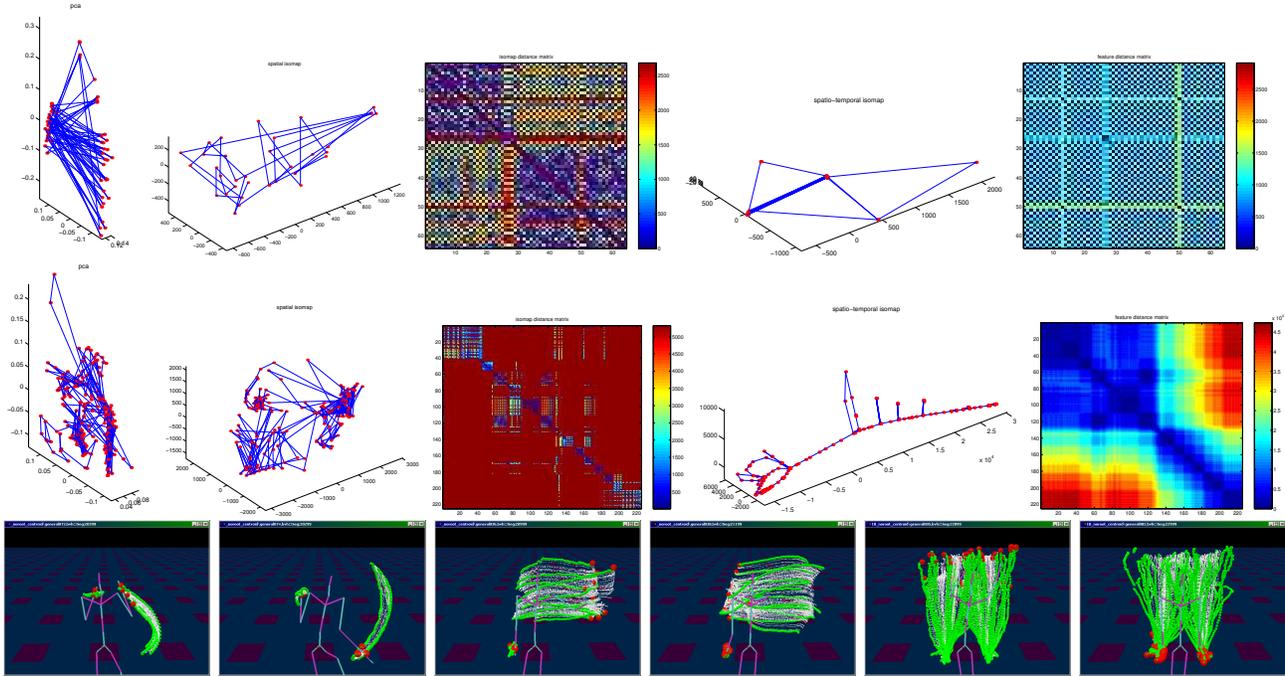
Figure 5. 3D embeddings of segmented reaching (top row) and multi-activity (middle row) human motion data from PCA (left), spatial Isomap (using 7 K-NN) with distance matrix (middle), and spatio-temporal Isomap (using 4 K-NN and $c_{textttCTN} = 100$) with distance matrix (right). Data points in each embedding are shown in red, with sequentially adjacent points connected with a blue line. (Bottom row) Extracted motion clusters generalized into manifolds.

ST-Isomap extracted 78 clusters from Input Motion 1 (multi-activity), including structurally significant clusters representing underlying behaviors and underrepresented transitions between these behaviors. Motion represented in selected clusters (and generalized to manifolds (Jenkins, 2003)) are shown in Figure 5. Noticable aspects in the ST-Isomap embedding are the flattening of the data into a near 1-manifold structure and loops sprouting from this manifold. The primary loop on the far left side of this manifold is due to the "horizonal arm waving" activity. This activity is performed during two non-adjacent intervals (segments 60-75 and 105-120), shown by the dark off-diagonal block in the distance matrix. By corresponding these segments of arm waving, the manifold creates a large loop consisting of behaviors that occur between the separate performances of arm waving. The smaller loops toward the middle of the manifold consisting of activities that alternate between a primitive-level behaviors, such as "punch outward" and "return to fighting posture".

5. Discussion

As with Isomap, ST-Isomap requires manual selection of various parameters, such as neighborhood se-

lection and embedding dimensionality. For local neighborhoods, ST-Isomap additionally requires a means for spatio-temporal correspondence within a neighborhood. Our experience has been that larger local neighborhoods are more difficult for performing accurate correspondences due to the greater number of points that are not spatio-temporally similar. Smaller neighborhoods provides easier correspondences that can be propagated via shortest-path computaton. Distal correspondence for such neighborhoods, however, requires that sampling of the underlying spatio-temporal process is dense enough such that split CTN components result. Split CTN components are multiple components that are representative of one underlying component. In contrast to underestimation of neighborhood size, our tendency is to overestimate embedding dimensionality. By overestimating dimensionality, we allow for all intra-CTN distances to be small, assuming additional dimensions provide better preservation of pairwise distances.

ST-Isomap in its current form provides a means to uncover spatio-temporal structure. In order to process larger data sets, however, we must either consider only a subset of the data (as landmarks) or bias the system through interval segmentation. To avoid subset land-

marks and heuristic biasing, a line of future research we plan to explore is uncovering spatio-temporal structure without necessarily computing all-pairs shortest-paths or embedding a full distance matrix.

6. Conclusion

We have presented ST-Isomap, an extension of Isomap nonlinear dimension reduction for data with both spatial and temporal relationships. Two instantiations of ST-Isomap were described for uncovering spatio-temporal processes in continuous and segmented data. Implementations of ST-Isomap were successfully applied to human motion and robot teleoperation data. As future work, we aim to perform comparisons of ST-Isomap with other unsupervised methods for segmenting and classifying data, such as methods based on time-series motifs (Chiu et al., 2003) and probability density function features (Kohlmorgen & Lemm, 2001).

Acknowledgments

References

Ambrose, R. O., Aldridge, H., Askew, R. S., Burridge, R. R., Bluethmann, W., Diftler, M., Lovchik, C., Magruder, D., & Rehnmark, F. (2000). Robonaut: Nasa's space humanoid. *IEEE Intelligent Systems*, *15*, 57–63.

Bengio, Y., Paiement, J.-F., & Vincent, P. (2003). Out-of-sample extensions for lle, isomap, mds, eigenmaps, and spectral clustering. *Advances in Neural Information Processing Systems 16 (NIPS 2003)*. Vancouver, British Columbia, Canada.

Brand, M. (2002). Charting a manifold. *Neural Information Processing Systems (NIPS 2002)*, *15*.

Chiu, B., Keogh, E., & Lonardi, S. (2003). Probabilistic discovery of time series motifs. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Washington, DC, USA.

Cohen, J. D., Lin, M. C., Manocha, D., & Ponamgi, M. K. (1995). I-COLLIDE: An interactive and exact collision detection system for large-scale envi-ronments. *Symposium on Interactive 3D Graphics* (pp. 189–196, 218). Monterey, California, USA.

Cox, T., & Cox, M. (1994). *Multidimensional scaling.* London: Chapman and Hall.

Howard, A., & Roy, N. (2003). The robotics data set repository (radish).

Jain, A., & Dubes, R. (1988). *Algorithms for clustering data.* Englewood Cliffs, NJ, USA: Prentice Hall.

Jenkins, O. C. (2003). *Data-driven derivation of skills for autonomous humanoid agents.* Doctoral dissertation, The University of Southern California.

Kohlmorgen, J., & Lemm, S. (2001). A dynamic hmm for on-line segmentation of sequential data. *Advances in Neural Information Processing Systems 14 (NIPS 2001)* (pp. 793–800). Vancouver, British Columbia, Canada: MIT Press.

Ng, A. Y., Jordan, M. I., & Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. *Advances in Neural Information Processing Systems 14 (NIPS 2001)* (pp. 849–856). Vancouver, British Columbia, Canada: MIT Press.

Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proc. of the IEEE*, *77*, 257–285.

Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, *290*, 2323–2326.

Schölkopf, B., Smola, A. J., & Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, *10*, 1299–1319.

Teh, Y. W., & Roweis, S. (2002). Automatic alignment of local representations. *Advances in Neural Information Processing Systems 15 (NIPS 2002)*. Vancouver, British Columbia, Canada.

Tenenbaum, J. B., de Silva, V., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, *290*, 2319–2323.

Varsta, M., Heikkonen, J., Lampinen, J., & del R. Milln, J. (2001). Neural processing letters. *Temporal Kohonen Map and the Recurrent Self-Organizing Map: Analytical and Experimental Comparison*, *13*, 237–251.

Zhou, D., Weston, J., Gretton, A., Bousquet, O., & Schölkopf, B. (2003). Ranking on data manifolds. *Advances in Neural Information Processing Systems 16 (NIPS 2003)*. Vancouver, British Columbia, Canada.