LOCALIZED INTERPOLATION FOR GRAPH SIGNALS

A. Mazarguil[1,2], *L. Oudre*[3], *N. Vayatis* [1,2]

[1] Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, F-94235, Cachan, France.
[2] Université de Paris, CNRS, Centre Borelli, F-75005, Paris, France.
[3] Université Sorbonne Paris Nord, L2TI, UR 3043, F-93430, Villetaneuse, France.

ABSTRACT

Graph structure is a powerful framework for the characterization of interaction between entities. In the context of graph signal processing, interpolation methods are based on the smoothness assumption that assumes that neighboring nodes present similar values. However, in several contexts, the strong relationship between two connected nodes may express different behaviour. In this paper, we propose a graph signal interpolation algorithm that uses a graph localization penalization on the reconstruction weights. These weights are learned from the data, thus allowing the use of signal anticorrelation on connected nodes in order to perform a more robust interpolation. The results displayed in the paper show that our approach is relevant when dealing with real data, for both smooth and non-smooth signals.

Index Terms— Graph Signal Processing, interpolation, missing data, imputation

1. INTRODUCTION

The presence of missing values in network data streams is an important issue that may arise in most real world scenarios. This loss of information can be due to many different factors, such as sensor default, loss in data transfer or system maintenance. Data interpolation methods are developed in this context, in order to replace the missing values with reliable estimates. Recent works have proposed to use the underlying structure of the data to guide the interpolation process. In particular, Graph Signal Processing (GSP) [1] has been introduced as an intuitive framework to deal with data lying on a network structure. Reconsidering the interpolation task in the context of graph processing tools appears as a natural approach to tackle the complexity of missing data handling.

The standard approach for interpolating missing graph signal values consists in basing the reconstruction under a smoothness assumption. In the GSP literature [1], the smoothness quantifies signal differences on connected nodes, weighted by the strength of the link between those nodes. Therefore, the underlying hypothesis is that strongly connected nodes carry similar values. Smooth graph interpolation techniques have been developed based on different

approaches, among them eigenspace projection [2], frame theory [3], kernel methods [4], graph matrix completion [5] and sparse methods [6].

Unfortunately, in several practical use cases, the smoothness hypothesis is not satisfied. For instance, strong negative correlations between neighboring nodes may be useful for reconstruction. This information cannot be used along the smoothness constraint, that will assume that neighboring nodes must have similar values. Another example is the case of information senders and receivers connected by a network (see Figure 1). As their role is inherently different, there is no a-priori reason for connected agents to exchange similar amounts of messages. Nevertheless the interaction count of an agent can be estimated from its neighbors. Finally, when considering multimodal graph signals (for instance meteorological data such as temperature/pressure), it seems natural that the measured quantities carry information on one-another, that however do not manifest as similar values.

Common approaches for relaxing the smoothness hypothesis consists in modifying the existing graph and seeking a new structure on which the initially non-smooth signal respects smoothness criterion [2, 7]. However, this approach does not take into account the original information carried by the graph structure, therefore possibly leading to deteriorated predictions.

In this paper, we propose a new approach for relaxing the smoothness assumption in graph signal interpolation. Our main assumption is that the proximity of two data sources in the graph structure indicates that they have a higher probability to share information, and therefore are more likely to contribute to the reconstruction of one another. However, as opposed to existing approaches, we will not assume that this shared information translates to similar graph values.

2. LOCALIZATION PENALTIES FOR GRAPH SIGNAL INTERPOLATION

Let us consider a weighted undirected connected graph $\mathcal{G} = (V, E, \mathbf{W})$ composed of a finite collection of n vertices $V = \{v_1, ..., v_n\}$, a set of edges $E \subset V \times V$ and a weight matrix $\mathbf{W} \in \mathbb{R}_+^{n \times n}$ for which the element $\mathbf{W}[i,j]$ quantifies the affinity between the nodes v_i and v_j. We assume that d
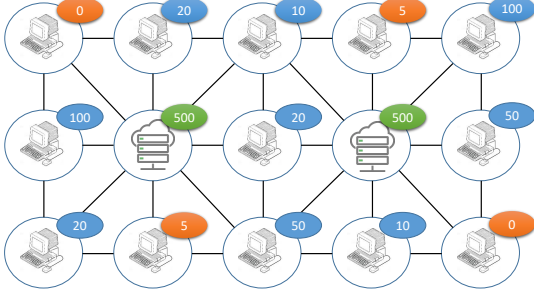
Fig. 1: Example of non-smooth graph signal. Here, the graph nodes present asymmetrical roles, either server or user. Therefore, the signal obtained from the interaction count is non-smooth, but the network carries useful information for an interpolation task.

(resp. p) nodes in V are defective (resp. non-defective) (with $n = d + p$) and denote V_d and V_p respectively the set of defective/non-defective nodes.

We consider two sets of graph signals lying on this graph $\mathcal{G}$.

- The first one, denoted as $\mathbf{X} \in \mathbb{R}^{n \times \ell}$ is composed of ℓ fully known graph signals. By rearranging its rows, matrix $\mathbf{X}$ can be rewritten as

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_d \\ \mathbf{X}_p \end{bmatrix}$$

where $\mathbf{X}_d \in \mathbb{R}^{d \times \ell}$ (resp. $\mathbf{X}_p \in \mathbb{R}^{p \times \ell}$) correspond to the restriction of matrix $\mathbf{X}$ to the defective (resp. non defective) nodes.

- The second one, denoted as $\mathbf{Y} \in \mathbb{R}^{n \times m}$ is composed of m graph signals for which the signal values on the defective nodes V_d are missing. With similar notations, $\mathbf{Y}$ writes as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_d \\ \mathbf{Y}_p \end{bmatrix}.$$

The aim of the interpolation task is this context is to provide an estimate for the missing data $\mathbf{Y}_d$.

The proposed approach consists in using the fully known signals $\mathbf{X}$ to learn a reconstruction rule from $\mathbf{X}_p$ to $\mathbf{X}_d$, and to use this rule to reconstruct $\mathbf{Y}_d$ from $\mathbf{Y}_p$.

2.1. Interpolation model

In this paper, the interpolation model is based on a standard affine model. We assume that the samples on defective nodes can be linearly derived from those of non defective nodes. Given a graph signal $\mathbf{x} = \begin{bmatrix} \mathbf{x}_d \\ \mathbf{x}_p \end{bmatrix}$, this model writes as

$$\mathbf{x}_d = \mathbf{A}\mathbf{x}_p + \mathbf{b} + \mathbf{e}, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{d \times p}$, $\mathbf{b} \in \mathbb{R}^d$ are affine reconstruction weights and $\mathbf{e} \in \mathbb{R}^d$ is a zero mean Gaussian reconstruction error term. The coefficient $\mathbf{A}[i, j]$ corresponds to the contribution of node $j \in V_p$ to the reconstruction of node $i \in V_d$.

2.2. The GPLR algorithm

In this section, we describe the Graph Penalized Linear Regression (GPLR) method, which falls into two parts : the estimation of the parameters $\mathbf{A}, \mathbf{b}$, and the interpolation of the missing value.

Step 1 - Model training.
The estimation of parameters $\mathbf{A}$ and $\mathbf{b}$ is performed on the set of fully known signals $\mathbf{X}$ by solving the following optimisation problem

$$\hat{\mathbf{A}}, \ \hat{\mathbf{b}} = \underset{\mathbf{A}, \ \mathbf{b}}{\operatorname{argmin}} ||\mathbf{X}_d - \mathbf{A}\mathbf{X}_p - \mathbf{b}\mathbb{1}_\ell^T||_F^2 + \mu \operatorname{Loc}(\mathbf{A}) \tag{2}$$

where

- The first term is a least square error which drives data fitting on the complete signal/nodes according to (1).

- The penalty term is the localization constraint defined as

$$\operatorname{Loc}(\mathbf{A}) = \sum_{\substack{i \in V_d \\ j \in V_p}} d_{\mathcal{G}}^2(i, j) \, \mathbf{A}[i, j]^2, \tag{3}$$

where $d_{\mathcal{G}} : V \times V \longrightarrow \mathbb{R}^+$ is a distance function on graph $\mathcal{G}$. Two graph distances will be tested in this paper : 1) the geodesic distance [8] that penalizes the contribution from node i to node j according to the length of the shortest path between them. 2) The pairwise distance in the D-laplacian eigenmap [9] of $\mathcal{G}$, that favors node to node contribution within the same cluster.

Step 2 - Missing values interpolation.
By using the reconstruction weights $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$ learned on $\mathbf{X}$, an estimate of the missing values $\hat{\mathbf{Y}}_d$ are derived from the known values $\mathbf{Y}_d$ by using the interpolation model (1):

$$\hat{\mathbf{Y}}_d = \hat{\mathbf{A}}\mathbf{Y}_p + \hat{\mathbf{b}}\mathbb{1}_m^T \tag{4}$$

2.3. Resolution of the optimization problem

This problem (2) is convex and admits a closed form solution for each row $i \in V_d$ of the matrix $\hat{\mathbf{A}}$:

$$\hat{\mathbf{A}}[i,:] = \left[\left(\bar{\mathbf{X}}_p \bar{\mathbf{X}}_p^T + \mu \boldsymbol{\Delta}_i \right)^{-1} \bar{\mathbf{X}}_p \right] \bar{\mathbf{X}}_d[i,:] \tag{5}$$

$$\hat{\mathbf{b}}[i] = \frac{1}{\ell} \mathbb{1}_\ell^T \left(\mathbf{X}_d[i,:] - \mathbf{X}_p^T \hat{\mathbf{A}}[i,:] \right) \tag{6}$$

with

- $\bar{\mathbf{X}}_p = \mathbf{X}_p - \frac{1}{\ell}\mathbf{X}_p \mathbb{1}_\ell \mathbb{1}_\ell^T$ the centered data matrix on defective nodes

- $\bar{\mathbf{X}}_d = \mathbf{X}_d - \frac{1}{\ell}\mathbf{X}_d \mathbb{1}_\ell \mathbb{1}_\ell^T$ the centered data matrix on non-defective nodes

- $\mathbf{\Delta}_i = \mathbf{Diag}\left(\left\{d_G(i,j)\right\}_{j \in V_p}\right) \in \mathbb{R}^{p \times p}$

The computational complexity of the closed form solution and derived approximated methods will be discussed in section **2.4**.

2.4. Additional remarks

Localization constraint. In model (1), the term $\mathbf{A}[i,j]$ quantifies the importance of node $j \in V_p$ in the reconstruction of node $i \in V_d$. The constraint introduced in (3) penalizes the use of *far away* nodes (in the sense of the graph structure) to be used in the interpolation process, hence enforcing the localization of the reconstruction. Moreover, the use of this constraint as a penalty allows a natural trade-off between quality of the reconstruction and distance between node pairs, allowing the model to seek distant contributions if the interpolation quality is improved. However, the resulting regression does not operate a strict node selection, meaning that a-priori, all entries of $\mathbf{A}$ will differ from zero. The notion of localisation is not carried by the support of the reconstruction weights, but by the proximity of large weight intensities, similarly to the notion of graph signal spread [10].

Affine model. In most smooth interpolation methods [2, 3, 4], the interpolated signal can be expressed as a matrix product of the known signal. The affine model proposed in (1) can be seen as a generalization of these approaches. This model is also compatible with standard assumptions in GSP such as bandlimitedness [1, 2] or low rank structure [5], that imply the existence of linear relations between node values.

Choice of the graph distance. Multiple graph distances can be used in the penalization term (3). A natural choice for the distance d_G is the geodesic distance. However, while being computationally expensive, the issue of generating this distance in the context of affinity graphs is complex and unsolved [11]. Our proposition is to perform the D-laplacian eigenmaps of the graph, and compute the pairwise distance in the resulting embedding. This approach is computationally cheaper for a reasonable value of the dimension D, and the use of this distance as a notion of proximity between graph nodes has been widely justified by the spectral clustering method [9]. Other graph based distances can be used, depending of the affinity notion that is relevant to the studied process, like the average node-to-node commutation time in random walk [12].

Complexity. The expression of the closed form solution requires d inversion of $p \times p$ matrices, for an overall complexity of $\mathcal{O}(d\,p^3)$, which can be relatively hard when working on large graphs. However, by using a method derived from

frame theory [3], an approximated solution can be deduced from successive iterations of the following operator :

$$f_\alpha(u) = \left(I - \alpha\left(\bar{\mathbf{X}}_p \bar{\mathbf{X}}_p^T + \mu\mathbf{\Delta}_i\right)\right)u + \alpha\bar{\mathbf{X}}_p\bar{\mathbf{X}}_d[i,:]$$

With the assumption that the graph distance is well defined ($\forall i \neq j, d_G(i,j) \neq 0$), the matrix $\bar{\mathbf{X}}_p \bar{\mathbf{X}}_p^T + \mu\mathbf{\Delta}_i$ is symmetric, defined positive with λ^+ (resp. λ^-) its highest (resp. lowest) eigenvalue. By setting $\alpha^* = \frac{2}{\lambda^+ + \lambda^-}$, the operator f_{α^*} is K-contractant with $K = \frac{\lambda^+ - \lambda^-}{\lambda^+ + \lambda^-} < 1$.

Therefore, any sequence of the form

$$\begin{cases} u_0 \in \mathbb{R}^p \\ u_{n+1} = f_{\alpha^*}(u_n) \end{cases} \tag{7}$$

converges exponentially fast to a fixed point of f_{α^*}, which is by design the solution of the problem (5). This approach reduces the overall complexity to $\mathcal{O}(d\,p^2 \ln(\epsilon^{-1}))$ where ϵ is the precision of the approximated solution.

3. EXPERIMENTS AND RESULTS

This section presents experimental results which support the use of the localization penalty introduced in (3) followed by a comparison with several standard interpolation methods on synthetic and real databases.

3.1. First validation of the localization penalty

In order to validate the intuition on the penalization term (3), let us consider a point cloud of $n = 200$ points uniformly sampled in a square, and the affinity graph obtained from this point cloud. We consider a set of $\ell = 50$ bandlimited graph signals on this graph (bandwidth equal to 50). All nodes are supposed to be non defective except for one and 50% of its values are removed. Figure 2 displays the reconstruction weights $\mathbf{A}^*$ learnt from (2) for different values of parameter μ that quantifies the penalization strength (1).

As the value of the parameter μ increases, the interpolation weights get more concentrated around the reconstructed node. Therefore, the proposed localization penalization seems relevant to enforce graph proximity between a defective node an the ones contributing to its reconstruction.

3.2. Evaluation protocol and data sets used

We now compare our interpolation method against state-of-the-art methods on different datasets:

- **Smooth.** A set of $\ell = 20$ band-limited signal on a random binomial connected graph of $n = 200$ nodes
- **Non smooth.** A set of $\ell = 20$ signals with 10 active low frequencies and 10 active high frequencies on a binomial connected graph of $n = 200$ nodes.

Method	Smooth		Non-smooth		Brittany		Mocap	
Missing ratio	0.3	0.5	0.3	0.5	0.3	0.3	0.3	0.5
Mean Interpolation	2.54 e-1	2.15 e-1	2.53 e-1	2.83 e-1	2.93 e+0	2.97 e+0	7.34 e+2	7.94 e+2
Narang et al. (2013) [2]	**7.27 e-16**	**6.92 e-16**	2.08 e-1	2.30 e-1	3.87 e+1	3.88 e+1	1.58 e+2	1.64 e+2
Chen et al. (2014) [5]	1.59 e-3	1.54 e-3	2.17 e-1	2.42 e-1	1.68 e+1	1.69 e+1	1.46 e+2	1.49 e+2
Linear Regression	**5.09 e-16**	**9.38 e-16**	1.51 e-1	2.06 e-1	5.53 e-1	5.99 e-1	2.15 e+0	9.12 e+0
Ridge Regression	4.74 e-7	2.71 e-6	1.51 e-1	2.05 e-1	5.47 e-1	5.90 e-1	2.19 e+0	7.22 e+0
Geodesic-GPLR	7.88 e-6	9.27 e-5	**4.42 e-2**	**1.73 e-1**	**5.28 e-1**	5.75 e-1	2.09 e+0	6.14 e+0
Embedding-GPLR	4.78 e-8	3.29 e-7	1.10 e-1	1.82 e-1	5.29 e-1	**5.74 e-1**	**2.05 e+0**	**5.51 e+0**

Table 1: RMSE for interpolation methods including GPLR proposed in this paper with two different choices of distance.
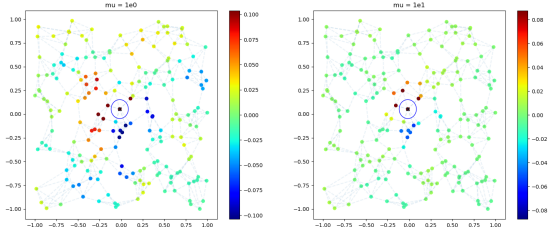
Fig. 2: Learned weights reconstructing the center node (circled) for $\mu = 1$ (left) and $\mu = 10$ (right) on the same dataset.

- **Brittany.** [13] The temperature of $n = 31$ locations in Brittany, each hour for a month ($\ell = 744$). The associated graph is computed by KNN ($K = 5$) with the sensors geographic positions.
- **Mocap.** [14] : A 3D times series dataset of $n = 44$ sensors fixed on a human body while performing tasks ($\ell = 7965$). The associated graph is obtained by linking neighbors sensors on the body and weighted by averaged sensor 3D distance. The multivariation of the signal is handled by considering a product of the sensor graph with a 3-cycle of weights 1.

For all the datasets, we randomly selected 25% of the nodes, and removed 30% or 50% of their values. The values of the parameter μ are selected by cross-validation. Several methods are tested and compared :

- Three baseline methods that do not take the graph structure into account: Mean Interpolation (each missing value is replaced by the mean of known values for the same sensor), Linear Regression (resolution of Problem (2) with $\mu = 0$) and Ridge Regression (resolution of Problem (2) where we assume that $d_{\mathcal{G}} = \mathbb{1}$).
- Two graph signal interpolation methods based on smoothness: Narang et al. 2013 [2] (Low band graph signal interpolation) and Chen et al. 2014 [5](Smooth graph matrix completion)
- Two variants of our interpolation method: Geodesic-GPLR ($d_{\mathcal{G}}$ is the geodesic distance) and Embedding-GPLR ($d_{\mathcal{G}}$ is the distance in the embedding space).

3.3. Results

The results are presented in Table 1. For each method and test case, we computed the root mean squared error of the reconstructions. In particular, GPLR displays top performance First of all, one can see that our methods perform quite well on the presented datasets. In particular, GPLR displays top performance on real world data, against non-regularized and smooth methods.

Smooth. As expected, the smooth signals are exactly reconstructed by the method of Narang et al. (2013) [2], which is designed for these signals. Because of the existence of the linear relation that exactly reconstructs the missing data, the non-penalized linear regression is also able to recover the same weights (or equivalents) without biais induced by regularization. Yet, we see that our method performs quite well even for smooth signals interpolation. Indeed, the smoothness criterion induces the existence of a good affine relationship with weight vectors localized on close hops.

Non smooth. For the non-smooth signals, our approach obtains the best results. The penalization term is a function of the weights intensities and is independent of the signs. This allows the regression model to tackle asymmetrical relationships between nodes, which naturally occurs in high band signals. The activity in high frequencies induces possible negative correlations between neighbors, that cannot be tackled approaches relying on a smoothness assumption.

Brittany. Regarding the temperature graph, the observed signal is smooth in the 3D space given the fact that it corresponds to measurements of a continuous physical variable. However, the smoothness characteristic drops when the graph is obtained from an irregular or non-exhaustive sample of the original space. In this case, the constructed graph can be an approximation of the underlying space, yet present relevant information on the studied process. By learning the reconstruction weights from the joint information of the graph structure and the known signals, this lack of correlation can be retrieved, focusing the reconstruction weights on relevant nodes. In other words, the learned data model is less sensitive to the graph construction technique that generated $\mathcal{G}$. The loss function being expressed as a sum of reconstruction error and localization, a natural trade-off operates, allowing the

model to seek information at more distant nodes if it implies an improvement of the reconstruction quality.

Mocap. Our approach obtains the best results. Thanks to the non-smooth framework, the multivariate setting is handled correctly by our method. The signals of each of the 3 dimensions (respectively X, Y and Z) are deeply correlated, therefore the values on the different components of each node carry information on other dimension even though these values may lie in different ranges.

3.4. Discussion

Influence of the missing ratio. Differences between our constrained approach and the standard ridge regression are especially visible when the percentage of missing data becomes large: in this case, the information carried by the graph is an addition to the quality of the reconstruction. Otherwise, results are almost similar since the locality of the weights can be retrieved in the data correlations without the prior knowledge of the graph structure.

Influence of the graph distance $d_{\mathcal{G}}$. The experiments tend to show that the proposed graph distances lead to interpolations of even quality. Our understanding of this behaviour falls into two parts. First, the notion of pairwise node proximity along the graph is preserved by both distances, therefore they induce a relatively similar penalization. On the other hand, as presented in discussion on the **brittany** results, a trade-off operates between the node to node distance and the quality of the reconstruction, leading to robustness of the model estimation regarding small variations of the distance function. Therefore, the choice of the suited graph distance is left to the user, depending on which specific characteristic should be enlightened by the graph structure (cluster closeness, isolation, optimal travel time, average travel time ...).

4. CONCLUSION

In this paper, we presented a new method for the interpolation of graph signals. The key assets of the the proposed method are the use of correlations / anti-correlations on close nodes, and the selection of relevant nodes along with a locality criterion. Therefore, our approach explores a wider class of relations than classical graph signal models, allowing the interpolation of non-smooth phenomenon on sensor networks.

5. REFERENCES

[1] D. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in *Proceedings of the IEEE Signal Processing Magazine*, vol. 30, 2012.

[2] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 5445–5449.

[3] X. Wang, P. Liu, and Y. Gu, "Local-set-based graph signal reconstruction," *Proceedings of the IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2432–2444, 2015.

[4] D. Romero, M. Ma, and G. B. Giannakis, "Kernel-based reconstruction of graph signals," *Proceedings of the IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 764–778, Feb 2017.

[5] S. Chen, A. Sandryhaila, J. Moura, and J. Kovacevic, "Signal recovery on graphs: Variation minimization," *Proceedings of the IEEE Transactions on Signal Processing*, vol. 63, 09 2015.

[6] A. Heimowitz and Y. C. Eldar, "Smooth graph signal interpolation for big data," 2018.

[7] A. E. C. Redondi, "Radio map interpolation using graph signal processing," *IEEE Communications Letters*, vol. 22, no. 1, pp. 153–156, Jan 2018.

[8] C. Godsil and G. Royle, "Algebraic graph theory," 2001.

[9] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Advances in neural information processing systems*, 2002, pp. 585–591.

[10] A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *Proceedings of the IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4338–4356, 2013.

[11] B. Pasdeloup, R. Alami, V. Gripon, and M. Rabbat, "Toward an uncertainty principle for weighted graphs," in *Proceedings of the 23rd European Signal Processing Conference (EUSIPCO)*. IEEE, 2015, pp. 1496–1500.

[12] T. Bonald, A. Hollocou, and M. Lelarge, "Weighted spectral embedding of graphs," in *Allerton Conference on Communication, Control, and Computing*, Oct 2018, pp. 494–501.

[13] B. Pasdeloup, V. Gripon, G. Mercier, D. Pastor, and M. G. Rabbat, "Characterization and inference of graph diffusion processes from observations of stationary signals," *Proceedings of the IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 3, pp. 481–496, Sep. 2018.

[14] M. Müller, T. Röder, M. Clausen, B. Eberhardt, B. Krüger, and A. Weber, "Documentation mocap database hdm05," 2007, http://mocap.cs.cmu.edu/.