Video Subject Inpainting: A Posture-Based Method

Abbas Koochari, and Mohsen Soryani

Abstract
Despite recent advances in video inpainting techniques, reconstructing large missing regions of a moving subject while its scale changes remains an elusive goal. In this paper, we have introduced a scale-change invariant method for large missing regions to tackle this problem. Using this framework, first the moving foreground is separated from the background and its scale is equalized. Then, a nonlinear dimension reduction is performed using manifold learning. We interpolate the missing values using motion analysis based on 3D body pose changing. The motion of the subject is considered as two different time series for which values are forecasted in the missing region. Since the subject motion may be non-periodic and forecasted series being different, a single series in the missing region is interpolated using blending of two forecasted series. Finally, the interpolated points are refined using weighted signal matching. In this method, a radial basis function (RBF) network has been used for mapping from spatial to manifold space. Experimental results, for a number of videos, show that the proposed approach is better than other methods and 3D body pose changing is smooth in the inpainted sequences.

Keywords**—**Video Inpainting, Manifold Learning, Time Series Forecasting, Body Pose.

I. INTRODUCTION

Filling-in and reconstruction of missing regions in image or video is known as inpainting or completion. These missing regions may be occurring haphazardly or manually. The main focus of inpainting is reconstructing these missing regions so that they are not detectable and are visually pleasant. There is a large variety of applications in this field of image and video processing [1-7]. Since spatio-temporal consistency must be preserved in video inpainting, it has more challenges than image inpainting. In this paper, a new video inpainting method based on motion analysis in manifold space has been proposed. This method analyses and forecasts subject motion considering the motion as a time series.

Manuscript received May 31, 2014; revised October 21, 2014; accepted February 24, 2015.
A. Koochari, Computer Engineering Department, Iran University of Science and Technology, Tehran, Iran (email: koochari@iust.ac.ir).
M. Soryani, Computer Engineering Department, Iran University of Science and Technology, Tehran, Iran (e-mail: soryani@iust.ac.ir).

A. Related Work

Video inpainting problem can be considered as image inpainting in 3 dimensions in which the third dimension is time and the main difference between these two types of inpaintings is preservation of spatio-temporal coherence. Several works have been done on image inpainting which are classified into two groups: Partial Difference Equation (PDE) and Exemplar-Based methods.

The PDE-based method works well for small regions without texture structure, and the Exemplar-based method works similar to texture synthesis techniques [8]. It is notable that applying these image-based algorithms to all video frames does not produce desirable results on video sequences. Bertalmio *et al.* in [9] applied a PDE-based method to all video frames. This approach is acceptable for small regions but for large regions causes blurring. Wexler *et al.* proposed a space-time completion method by modeling the problem to a local and global optimization problem [10-11]. Since this approach performs an exhaustive search to find a cubic patch, it has huge computations. Also, it is acceptable for periodic subjects motion without scale changing.

Zhang *et al.* used motion segmentation and isolated different layers of a video, and then performed motion compensation for filling in the missing region [12]. However, obtaining different video motion layers is not so accurate. Patwardhan *et al.* separated the moving foreground from the background, and then reconstructed the missing data separately [13].The approach proposed by Patwardhan does not consider the scale changing in subjects and does not work well when the moving subject faces a large missing region. Shen *et al.* in [14] proposed a method in which by constructing a space-time manifolds volume and then rectifying the subject volume, the perspective distortion is repaired. Their proposed method produces distortion and is suitable for small missing regions. Most proposed methods are not acceptable for completing large missing regions and cannot synthesize a new subject.

A subject-based method has been proposed in [15-16] in which sets of subject templates are extracted and used for reconstructing damaged subjects in corrupted frames. The drawback of this method is jumping and body pose change when a subject enters and exits from an occluding area. Also, this method creates anomaly samples. Koochari *et al.* [17]

segmented the moving subject from background and constructed a large mosaic from subject frames. Then, they inpainted the mosaic by applying a large exemplar-based method. This method is suitable for periodic motions.

In [18] a Rank minimization method in manifold space for reconstruction of missing regions is proposed. This method reshapes the subject descriptors to a Hankel matrix and solves it using matrix completion approach. This approach has only been examined on small missing regions and subject scale changing has not been considered in it.

In our proposed method two challenges of large missing regions and subject scale changing are considered and they are solved by rectification of the subjects and interpolation of the missing values using motion analysis based on 3D body pose change in manifold space.

B. *Overview of Our Work*

The primary contribution of this paper is the forecasting of time series for interpolation of missing coefficients in low dimensional manifold and smooth transition between frames by 3D body pose consideration. In the proposed method, after separating the moving foreground from background the following steps are performed for moving subject inpainting:

1. Rectification of moving subjects: with respect to scale changing of moving subjects (in test videos), rectification (equalizing subject scale) is done by constructing a small mosaic image.
2. Mapping from spatial to manifold space: the input subject is mapped to the manifold space and 3 Locally Linear Embedding (LLE) [19] coefficients of each input are maintained.
3. Time series forecasting and missing coefficients interpolation: in this step corrupted moving subject coefficients in low dimensional manifold are obtained by time series forecasting, and then interpolation is done.
4. Weighted signal matching and inverse mapping from manifold space to input: finally, a weighted signal matching is applied to the obtained coefficients, and then a Radial Basis Function (RBF) [20-21] is used for inverse mapping from manifold to input space.

Schematic overview of the moving subject inpainting method is shown in Figure 1.

Our proposed method is explained in Section 2 in details. Section 3 presents experimental results and Section 4 concludes the paper with a discussion on future works.

II. The Proposed Algorithm

In this section the proposed method is explained in details. At first, the moving foreground is separated from the background. Moving subject detection usually is done by background subtraction. For this purpose, an accurate estimation from background is obtained and the moving foreground is detected by subtracting each frame from it.

Several methods have been proposed for background modeling. In this paper, a Gaussian Mixture Model (GMM)-based method proposed in [22] has been used. Afterwards the detected subjects are prepared for the next step where moving subjects and background are inpainted separately.

A. *Moving Subject Inpainting*

Steps of moving subject inpainting are as follows:

A.1. Scale equalization of moving subjects

In some of the selected videos subject's scale changes and therefore equalization of subject's scale is essential in all frames. Periodic duration of subject's motion can be calculated to obtain the amount of projective distortion. Another way to calculate the amount of projective distortion, as depicted in Figure 2, is line fitting in which the maximum and minimum points of subjects, when feet of the subject are close, are connected together. Also, if camera is not orthogonal to the moving subject, the head and feet of the subject can be obtained by Eigen analysis. After calculating the amount of projective distortion, subject scale is equalized in all frames by affine and metric rectification. To remove the projective distortion and recover the affine properties from a mosaic frame, the vanishing line, l_v, is mapped into the line at infinity [24]. This work is done by a homographic matrix as

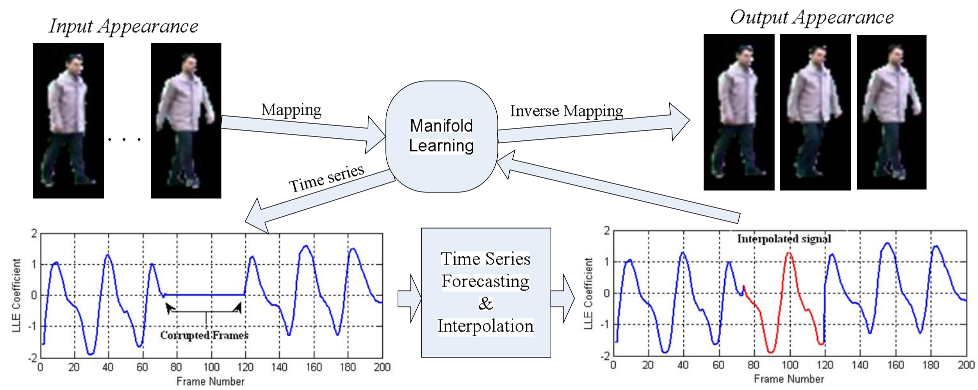

Fig. 1. Schematic overview of the proposed algorithm for moving subject inpainting.

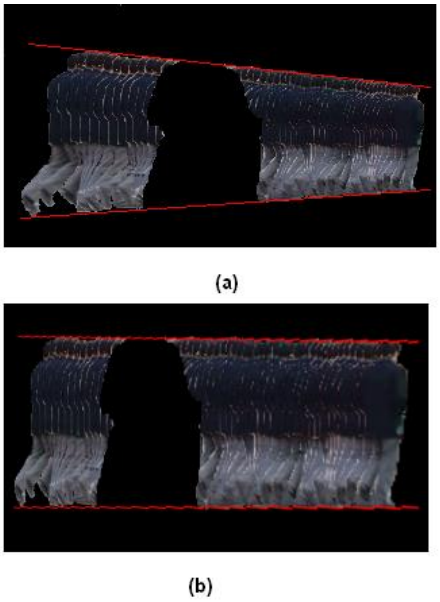

Fig. 2. Rectification and subject scale converting (Panels (a) and (b) are the small mosaic before and after elimination of projective distortion).

follows:

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix} \quad (1)$$

where $l_v = [l_1\ l_2\ l_3]^T$. Then, the H_a Matrix is obtained for affine to metric rectification using two sets of orthogonal lines. Here, Cholesky factorization is used to obtain the K matrix according to [24] (K is a 2×2matrix). The H_r matrix is obtained by composition of the affine and the metric transformations. Panels (a) and (b) in Fig. 2 show the small mosaic before and after applying the H_r matrix respectively. The H_r matrix is obtained by composition of the affine and the metric transformations.

$$H_a = \begin{bmatrix} K & 0 \\ 0 & 1 \end{bmatrix} \quad (2)$$

After calculating the H_r matrix, it is applied to all frames separately for equalization of the subject scale.

A.2. Mapping from spatial to manifold space

Since human appearance changes in different frames, dimension reduction is used for capturing the subjects appearance. Recently manifold learning is used for representation of images in low dimension. LLE is a nonlinear dimensionality reduction method which can represent local information by finding a set of nearest neighbors of each point [19]. The algorithm proposed by Elgammal *et al.* [25] learned human appearance using LLE and estimated the 3D pose from the appearance using a nonlinear mapping. Since human appearance can be represented by 3 coefficients in manifold space [21,25], moving subjects are mapped to manifold using LLE to capture local information. There are other similar global and local non-linear manifold learning methods such as ISOMAP [26], Laplacian Eigenmap [27], and Hessian LLE [28].

Let$I_t$bethe moving subject in frame t, t = [1, T], sequence of correct frames is denoted by x_t where $1 \le t < r$ and $s < t \le T$ and sequence of damaged frames is denoted by d_t where $r < t < s$. Each correct frame, x_t, is normalized to a w×h window and reshaped to a (w×h)×1 vector, and then mapped to a 3×1 vector, y_t, in a low dimensional manifold using LLE (w and h are width and height of normalized window, respectively).

Dimension reduction steps by LLE are as follows:

1. Find k nearest neighbors, x_j, for each x_i.
2. Find the weights w_{ij} which best reconstruct x_i from its neighbors and minimize the following equation

$$\sum_i \| x_i - \sum_{j=1}^{k} w_{ij} x_j \|^2$$

where $\sum_j w_{ij} = 1$.

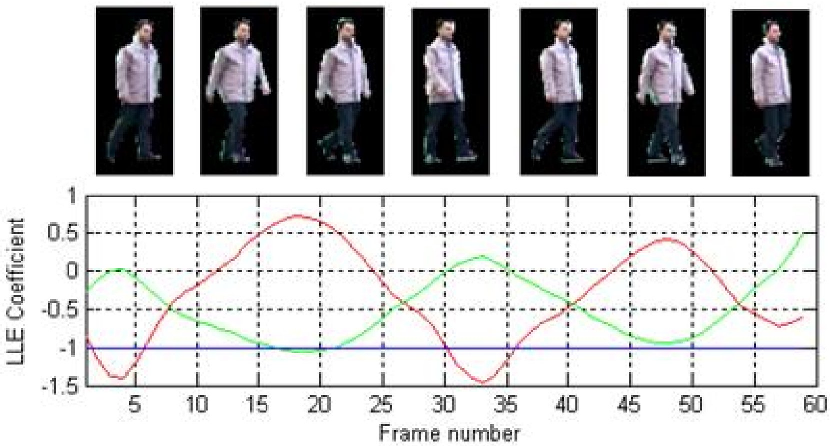

Fig. 3. Mapping from input to low dimensional manifold.

3. Calculate the vector y_i which minimizes the reconstruction error using the weights w_{ij}
$$\sum_i \| y_i - \sum_{j=1}^{k} w_{ij} y_j \|^2$$

After dimension reduction using manifolds, missing coefficients must be estimated for interpolation of corrupted data. Since the LLE coefficients of sequences are similar to time series, times series prediction are used for construction of the missing region. Figure 3 illustrates a sequence in spatial and manifold space.

The parameter k, which is the number of nearest neighbors, is important to obtain the LLE coefficients. A low number of neighbors causes over-fitting and a high number causes local information loss.

For interpolation of coefficients in the missing region, each signal (coefficient sequence) is considered as two different time series which move toward the corrupted area from start and end points of the signal.

A.3. Time series forecasting and missing coefficients interpolation

After mapping the uncorrupted subjects to the low dimensional manifold, each signal (a coefficient of frame sequences) is considered as two time series starting from begining and end points of the signal to the missing region. The reseaon for using time series is the semi-periodic behavior of the subject motion in manifold space. For automatic forecasting of time series an appropriate model must be specified and its parameters be estimated. Popular automatic methods for time series forecasting usually use one of the exponential smoothing or Autoregressive Integrated Moving Average (ARIMA) models. In this work a step-wise algorithm is used for forecasting using ARIMA model [31]. If X_t be time series data in time t, ARIMA (p,d,q) model is introduced as follows:

$$(1-\sum_{i=1}^{p}\Phi_i B^p)(1-B)^d X_t = (1+\sum_{i=1}^{q}\theta_i B^q)\varepsilon_t \quad (3)$$

In this model, B is the shift operator, ε_i is the time error, p is the order of autoregressive part, d is the non seasonal differences and q is the number of lagged forecast errors in the prediction equation, which is related to moving average part [31].

An important problem in automatic forecasting is the appropriate selection of ARIMA model order. The best model is obtained based on an information criterion named AIC (Ailake Information Criteria) with parameters p and q as follows:

$$AIC = -2\log(L) + 2(p+q) \quad (4)$$

where L is the maximized value of the likehood for the estimated model.

To obtain the best model, the d parameter is determined by examining the data on an ARIMA model with a root first. If an acceptable result was obtained, the derivatives of data is examined too. This procedure runs iteratively by incrementing the d parameter until no undesirable result is achieved. Since finding a model with best AIC, based on checking different values of p and q parameters, is very time consuming and computationally expensive, a step-wise algorithm is done as follows:

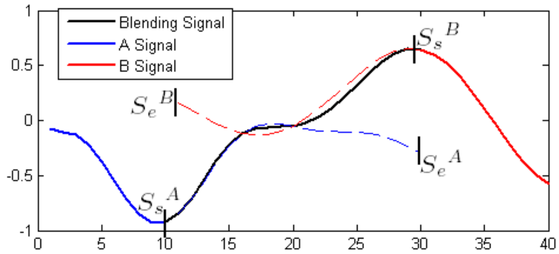

Fig. 4. Blending in junction region of two signals.

1. Data are tested with four models: ARIMA (0,d,0), ARIMA (0,d,1), ARIMA (1,d,1) ARIMA (2,d,2). Then the model with the lowest AIC is selected as the current model.
2. p and q parameters are changed by ±1 and the new models are evaluated again. If the value of AIC obtained from the new models is better than that of the current model, then the current model is replaced with the best model. This scheme is continued to get the best model so that none of the new obtained AIC's is better than the current model.

Considering that subject motion may not be periodic in some videos, the two forecasted time series necessarily are not similar and they do not meet each other in one point. Therefore after forecasting the time series in the missing region, the forecasted time series are blended in their junction point to preserve the continuity of two signals. Since the beginning section of the forecasted signal is more accurate than its terminal section, transition's duration is considered half of the missing region, so that signal blending is done from the middle point of the missing region between two time series as depicted in Figure 4.
where S_e^A,S_s^A,S_e^B and S_s^B are the begining and the end sections of signals A and B which are used in signals blending. A decreasing sigmoiedy function like $a = 0.5\cos(\pi b) + 0.5$ is used to combine these two signals, so that b is changed between 0 to 1. Signal A and B are multiplied by α and 1-α respectively.

A.4. Weighted signal matching and mapping from manifold space to input

In this step, since reconstructed samples must not be much different to input samples, signal matching is used to obtain the most similar reconstructed frames to input frames in manifold space. For this purpose, two different techniques are used: Partial matching and General matching.

In the Partial Matching, signal matching is done only for the beginning and end sections of time series in the missing region separately. Consequently, we will have a new synthesized sample in the junction point of two signals by blending. Therefore, new samples are synthesized for subjects who have non-periodic motion, but if the motion of the subject is periodic, new samples are not created.

In the general matching, the whole reconstructed signal is matched and replaced by the best candidate of correct signal. Due to the importance of signal in entrance and exit points of the missing region, weighted signal matching is done to have smooth body pose changing in these locations. For the weighted signal matching $w(x;\sigma,c)=1-\exp(\frac{-(x-c)^2}{2\sigma^2})$ is used as the weight and the best matching signal is obtained by minimizing the sum square error in the bounded region as in equation (5):

$$\min_{p\in t}\sum_{i=1}^{N}[(X_t(p+i)-S_I(i)).w(i)]^2 \qquad (5)$$

where N is the number of damaged frames, t is the total number of frames, SI is the interpolated signal and X_t is the LLE manifold of correct frames.

After weighted signal matching, inverse mapping to input is done using Radial Basis Functions which are trained using correct frames. Using the Radial Basis Functions, we can synthesize new samples.

B. Background Inpainting

Background inpainting is done separately using a greedy patch-based algorithm which proposed by Criminisi *et al.* [32]. In this algorithm, the missing region (the target region) is filled in by source region (the safe region). To fill the holes of the image, a square patch is considered for each pixel in boundary of the target region and then this square patch is filled in by the patch which has the most similarity to the source region. To keep the structure of the image, a priority value is calculated for each pixel in the boundary of target region, and the pixel with maximum priority is selected to fill in at first. The priority value of the pixel is the product of confidence and data terms as follows: $P(p)=C(p).D(p)$.

The confidence term C(p) is the number of undamaged pixels divided to all pixel of surrounding

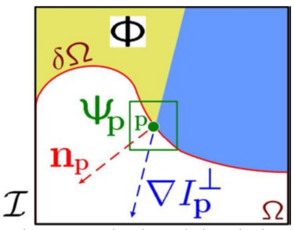

Fig. 5. Notations in Exemplar-based inpainting (borrowed from Criminisi et al. [32]).

patch of the pixel P. The data term D(p) is high if gradient of the pixel and orthogonal to boundary of the pixel are unidirectional which is explained as follows:

$$C(p)=\frac{\sum_{q\in\Psi_p\cap\bar{\Omega}}C(q)}{|\Psi_p|} \qquad (6)$$

$$D(p)=\frac{|\nabla I_p^{\perp}.n_p|}{\alpha} \qquad (7)$$

where Ψ_p is a patch centered in location p, $|\Psi_p|$ is the area of the patch, α is the factor for normalization (e.g. α=255 for grayscale images), n_p is normal to the boundary of the target region and $\nabla I_p^{\perp}$ is an isophote in location p (see Fig. 5 for notations). The target region, source region and entire image are denoted as Ω, Φ and I, respectively.

The patch $\Psi_{\hat{p}}$ with maximum priority is found (i.e., $\Psi_{\hat{p}} \mid \hat{p}=\arg\max_{p\in\partial\Omega} P(p)$) and the best matching patch in the source region with $\Psi_{\hat{p}}$ is selected and copied into $\Psi_{\hat{p}}$. At the end, the confidence terms for all pixels on the boundary between source and target regions are updated. This algorithm is run iteratively until the missing region fades away.

III. Experimental Results

To evaluate the proposed method, five video sequences were examined. Scale changing and facing with large missing regions is occurred in two of the test video sequences. In the first video which was used in [18], a person moves from right to left and passes behind another person. There are two methods for occluded subject inpainting. In the first method after obtaining the best matching subject, the corrupted parts of the subject are replaced by the corresponding parts of the matching subject. In the other method which is easier, the occluded subject is replaced entirely with the best matching subject. Since human is a deformable subject, replacing the entire subject is visually better than partial replacement. Fig. 6 shows results of the two mentioned methods.

As depicted in Fig. 6, replacing the entire subjectis more acceptable because in second row discontinuity is clearly visible. Note that in the above sequence motion of the subject is periodic and changes in human appearance are very low, nonetheless partial replacement does not produce good results. Therefore to preserve subject structure and to avoid of anomalous subject creation, in the rest of experiments the second method, i.e. entire subject replacement, is considered.

Scale changing is occurred in the second and the third video samples which were captured by a hand-held camera (see panels (a) and (b) in Figure 7). In the second video, a person crosses behind a statue (see Figure 11). The statue for this video was created manually. After subject scale equalization (by creating a small mosaic), the proposed method was applied. Figure 8 shows the results for a number of frames when the subject enters to and exits from the occluding area.

As shown in Figures 8(c) and 8(d) body pose of the subjects have been preserved. If partial replacement had been used, the abnormal results would have been created because subject's appearances are different in the best matching subject and the corrupted subject.

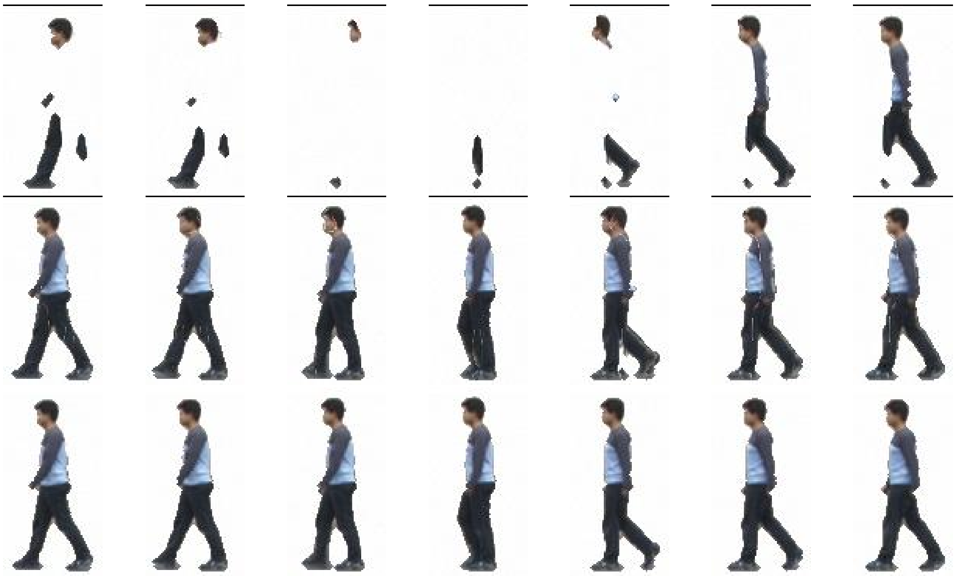

Fig. 6. Results of our proposed algorithm on the first video (First row: corrupted subjects from right to left, Second row: inpainted subjects obtained by replacement of corrupted parts with parts of the best forecasted subject, Third row: inpainted subjects obtained by replacement of entire corrupted subjects with the best forecasted subject).

(a) (b)

Fig. 7. Scale changing in input videos, (a) second video sample, and (b) Third video sample.

(a) (b)

(c) (d)

Fig. 8. Results of the proposed method for second video (Panels (a) and (b) show three consecutive subject frame when the subject enters to and exits from the missing region respectively. Panels (c) and (d) show the results the proposed method based on body pose changing).

In the third video a man passes behind a tree. After subject's scale equalization, the proposed method was applied. As can be seen in Fig. 8 when the subject exits from the occluding area, body pose changesare smooth. The first three images are parts are parts of the occluding area and therefore the whole subject has been replaced.

As Fig. 9 shows, the body pose is changed smoothly and the inpainted sequence is visually desirable. To compare the proposed method with other methods, the 'jumping girl' video, which was captured and used by Wexler *et al.* [11], has been used. In this video, a girl moves from left to right and passes behind an occluding subject (a person). Figure 10 shows results of the proposed method compared to the results of Wexler *et al.* (space time video completion) [11] and Venkatesh *et al.* [16] methods, respectively. In the first row of Figure 9 the occluding subject has been shown with a black mask.

As mentioned above, background inpainting is done separately using an exemplar-based method proposed by Criminisi*et al.* [32]. Figure 11 shows the reconstructed backgrounds of the tested videos in this research. In these images the statue, tree, black mask and board are inpainted from the background.

Fig. 9. Some frames of the original and the inpainted sequences for the third video (The first row shows consecutive frames when the man exits from behind the tree, and the second row shows the inpainted sequence using the proposed method).

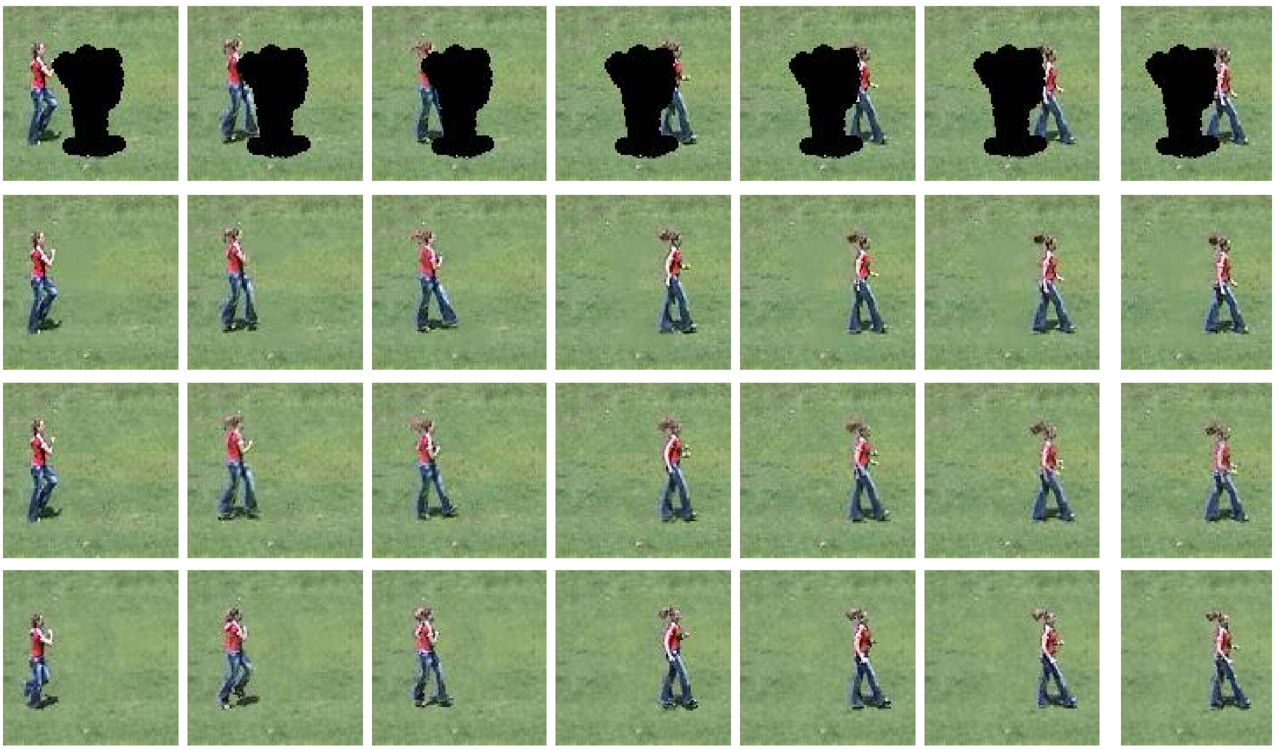

Fig. 10. Comparison of the results of our proposed method with the results of algorithms proposed by Wexler et al. [11] and Venkatesh et al. [16] (The first row shows the girl passing behind the occluding mask; the second row shows the inpainted sequence using space-time completion method [11]; the third row shows the inpainted sequence using Venkatesh's method [16] and finally the fourth row shows the results of our algorithm).

(a) (b) (c)

Fig. 11. Results of background inpainting for the tested videos (Panels (a) through (c) in the first row show backgrounds of the original sequences and panels (a) through (c) in the second row show backgrounds of the inpainted sequences).

Finally, the output video is obtained by composing the inpainted foreground and background frames. To improve the output, in the bordering area between foreground and background, a simple alpha matting method is applied. The resulting videos can be viewed at:

http://webpages.iust.ac.ir/koochari/proj/poseinpainting.html.

IV. Conclusion and Future Work

In this paper, a new video inpainting method has been proposed using manifold learning. In this method, the moving subjects are separated from the background, scale equalized and mapped to the LLE manifold. By preserving three LLE coefficients for each subject, subject body pose is derived. Afterwards, missing coefficients for occluding subject area in low dimensional manifold are interpolated using time series forecasting, blending and weighted signal matching. For interpolation of these coefficients, each signal is considered as two time series which move toward the occluding subject area. Then an ARIMA model is estimated automatically for forecasting each time series. Afterwards blending and weighted signal matching are applied to the interpolated signal to obtain most similar sequences between correct and interpolated parts. Radial Basis Functions (RBF), trained by correct moving subjects, are used to map from low dimensional manifold to input space. Finally, background inpainting is done using an exemplar-based method and the output video is obtained by composing the inpainted foreground and background frames. In this method, the subject motion based on body pose, which has been analyzed in manifold space, is used to interpolate the subject motion in the missing area. Also spatio-temporal continuity is considered using an ARIMA-based time series model. Results of the algorithm are very good for tested video sequences and subject pose changing is smooth in consecutive frames.

Future studies can be conducted for more challenging situations such as moving camera and non-periodic motion of subjects. Also subject representation and its motion analysis for synthesizing a new subject pose using different parts of uncorrupted subjects, is another interesting subject which can be considered in future works.

References

[1] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester,“Image Inpainting,”*In Proc. of the ACM SIGGRAPH Conference on Computer Graphics*, 2000, pp. 417-424.

[2] M. M. Oliveira, B. Bowen, R. McKenna, Y.-S. Chang, “Fast Digital Image Inpainting,” *In Proc. of the International Conf. on Visualization, Imaging and Image Processing* (VIIP), 2001, p. 261–266.

[3] M. Bertalmio, L. Vese, G. Sapiro, S. Osher, “Simultaneous structure and texture image inpainting,” *IEEE Transactions on Image Processing*, 2003,12, (8), pp.882–889.

[4] J. Sun, L. Yuan, J. Jia, H.Y. Shum, “Image Completion with Structure Propagation,” *ACM Transactions on Graphics* (TOG), 2005, 24, pp.861-868.

[5] H. T. Ho, R. Goecke, “Automatic parametrisation for an image completion method based on markov random fields,” *In Proc. IEEE Int. Conf. on Image Processing*, 2007, 3, pp. 541-544.

[6] D. Liu, X. Sun, F. Wu, S. Li, Y.Q. Zhang, “Image Compression With Edge-Based Inpainting,” *IEEE Transactions on Circuits and Systems for Video Technology,* 2007, 17, (10), pp.1273–1287.

[7] Y. Matsushita, E. Ofek, W. Ge, X. Tang, H.Y. Shum, “Full-Frame Video Stabilization with Motion Inpainting,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28, (7), pp.1150-1163.

[8] A. Efros and T. Leung, “Texture synthesis by non-parametric sampling,” *In Proc. of IEEE International Conference Computer Vision*, Corfu, Greece, September 1999, pp. 1033-1038.

[9] M. Bertalmio, A.L. Bertozzi, G. Sapiro, “Navier-stokes, fluid dynamics, and image and video inpainting,” *In Proc. IEEE Conference Computer Vision and Pattern Recognition* (CVPR), 2000, 1, pp.355–362.

[10] Y. Wexler, E. Shechtman, M. Irani, “Space-time video completion,” *In Proc. IEEE Computer Society Conference Computer Vision and Pattern Recognition* (CVPR), 2004, 1, pp.120–127.

[11] Y. Wexler, E. Shechtman, M. Irani, "Space-time completion of video," IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29, (3), pp. 463-476.

[12] Y. Zhang, J. Xiao, M. Shah, "Motion layer based object removal in videos," *In Proc. of IEEE Workshop on Applications of Computer Vision*, 2005, pp. 516–521.

[13] K. A. Patwardhan, G. Sapiro, M. Bertalmio, "Video inpainting under constrained camera motion," *IEEE Trans. on Image Processing*, 2007, 16 (2), pp. 545–553.

[14] Y. Shen, F. Lu, X. Cao, H. Foroosh, "Video Completion for Perspective Camera Under Constrained Motion," *In Proc. of the 18th International Conference on Pattern Recognition* (ICPR06), 2006, 3, pp. 63-66.

[15] S. Cheung, J. Zhao, M. V. Venkatesh, "Efficient Object-Based Video Inpainting," In Proc. *IEEE International Conference on Image Processing* (ICIP06), 2006, pp.705-708.

[16] M. V. Venkatesh, S. Cheung, J. Zhao, "Efficient object-based video inpainting," *Pattern Recognition Letters*, 2009, 30(2), pp. 168-179.

[17] A. Koochari and M. Soryani, "Exemplar-based video inpainting with large patches," *Journal of Zhejiang University Science C*, 2010, 11(4), pp. 270-277.

[18] T. Ding, M. Sznaier and O. Camps, "A Rank Minimization Approach to Video Inpainting," *IEEE International Conference on Computer Vision*, Rio de Janeiro, Brazil, October 2007.

[19] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," Sciene, 2000, 290(5500), pp.2323–2326.

[20] T. Poggio and F. Girosi, "Networks for approximation and learning," *In Proc. of IEEE*, 1990, vol. 78, pp.1481–1497.

[21] H. Lim, V. I. Morariu, O. I. Camps, and M. Sznaier, "Dynamic appearance modeling for human tracking," In Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR), 2006, 1, pp. 751–757.

[22] M. Piccardi, "Background subtraction techniques: a review," *In Proc. of IEEE SMC 2004 International Conference on Systems, Man and Cybernetics*, Oct. 2004, vol. 4, pp. 3099-3104.

[23] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," *In Proc. of the 17th Int. Conf. on Pattern Recognition*, 2004, 2:28- 31

[24] R. I. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, (Cambridge University Press, 2004).

[25] A. Elgammal and C. S. Lee, "Inferring 3d body pose from silhouettes using activity manifold learning," *IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), 2004, Vol. 2, pp. 681–688.

[26] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, 2000, vol. 290, pp.2319–2323.

[27] D. L. Donoho and C. E. Grimes, "Hessian eigenmaps: locally linear embedding techniques for high-dimensional data," *In Proc. of the National Academy of Arts and Sciences*, 2003, vol. 100, pp. 5591–5596.

[28] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, 2003, 15(6), pp.1373–1396.

[29] T.C. Mills, Time Series Techniques for Economists. *Cambridge University Press*, 1990.

[30] G. Box, G. M. Jenkins, and G. C. Reinsel. Time Series Analysis: Forecasting and Control, third edition. *Prentice-Hall*, 1994.

[31] R.J. Hyndman, Y. Khandakar, "Automatic Time Series Forecasting: The forecast Package for R," *Journal of Statistical software*, 2008, vol. 27, no. 3.

[32] A. Criminisi, P. Perez, K. Toyama, "Region filling and object removal by exemplar-based inpainting," *IEEE Trans. on Image Processing*, 2004, 13, (9), pp. 1200–1212.