Available online at www.sciencedirect.com
SCIENCE DIRECT®

Pattern Recognition 38 (2005) 1033–1043

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

Behavior classification by eigendecomposition of periodic motions

Roman Goldenberg*, Ron Kimmel, Ehud Rivlin, Michael Rudzsky

Computer Science Department, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel

Received 23 January 2004; received in revised form 25 October 2004; accepted 8 November 2004

Abstract

We show how periodic motions can be represented by a small number of eigenshapes that capture the whole dynamic mechanism of the motion. Spectral decomposition of a silhouette of a moving object serves as a basis for behavior classification by principle component analysis. The boundary contour of the walking dog, for example, is first computed efficiently and accurately. After normalization, the implicit representation of a sequence of silhouette contours given by their corresponding binary images, is used for generating eigenshapes for the given motion. Singular value decomposition produces these eigenshapes that are then used to analyze the sequence. We show examples of object as well as behavior classification based on the eigendecomposition of the binary silhouette sequence.

Keywords: Visual motion; Segmentation and tracking; Object recognition; Activity recognition; Non-rigid motion; Active contours; Periodicity analysis; Motion-based classification; Eigenshapes

1. Introduction

Futurism is a movement in art, music, and literature that began in Italy at about 1909 and marked especially by an effort to give formal expression to the dynamic energy and movement of mechanical processes. A typical example is the 'Dynamism of a Dog on a Leash' by Giacomo Balla, who lived during the years 1871–1958 in Italy, see Fig. 1 [1]. In this painting one could see how the artist captures in one still image the periodic walking motion of a dog on a leash. Following a similar philosophy, we show how periodic motions can be represented by a small number of eigenshapes that capture the whole dynamic mechanism of periodic motions. Singular value decomposition of a silhouette of an object serves as a basis for behavior classification by principle component analysis. Fig. 2 present a running horse video sequence and its eigenshape decomposition. One can see the similarity between the first eigenshapes—Fig. 2(c and d), and another futurism style painting "The Red Horseman" by Carlo Carra [1]—Fig. 2(e). The boundary contour of the moving non-rigid object is computed efficiently and accurately by the fast geodesic active contours [2]. After normalization, the implicit representation of a sequence of silhouette contours given by their corresponding binary images, is used for generating eigenshapes for the given motion. Singular value decomposition produces the eigenshapes that are used to analyze the sequence. We show examples of object as well as behavior classification based on the eigendecomposition of the sequence.

2. Related work

Motion-based recognition received a lot of attention in the field of image analysis. The main reason is that direct usage of temporal data can improve our ability to solve a

* Corresponding author. Tel.: +972 4 8294940; fax: +972 4 8293900.

E-mail addresses: romang@cs.technion.ac.il (R. Goldenberg), ron@cs.technion.ac.il (R. Kimmel), ehudr@cs.technion.ac.il (E. Rivlin), rudzsky@cs.technion.ac.il (M. Rudzsky).

0031-3203/$30.00 © 2005 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2004.11.024

number of basic computer vision problems. Examples are image segmentation, tracking, object classification, etc.

In general, when analyzing a moving object, one can use two main sources of information: changes of the moving object position (and orientation) in space, and the object's deformations.

Object position is an easy-to-get characteristic, applicable both for rigid and non-rigid bodies. It is provided by most of existing target detection and tracking systems. A number of techniques [3–6] were proposed for the detection of motion events and recognition of various types of motions based on the analysis of the moving object trajectory and its derivatives. Detecting object orientation is a more challenging problem. It is usually solved by fitting a model that may vary from a simple ellipsoid [6] to a complex 3-D vehicle model [7] or a specific aircraft-class model adapted for noisy radar images as in Ref. [8].

Fig. 1. 'Dynamism of a Dog on a Leash' 1912, by Giacomo Balla, Albright-Knox Art Gallery, Buffalo.

While object orientation characteristic is more applicable for rigid objects, it is object deformation that contains the most essential information about the nature of the non-rigid body motion. This is especially true for natural non-rigid objects in locomotion that exhibit substantial changes in their apparent view. In this case the motion itself is caused by these deformations, e.g. walking, running, hopping, crawling, flying, etc.

There exists a large number of papers dealing with the classification of moving non-rigid objects and their motions, based on their appearance. Lipton et al. describe a method for moving target classification based on their static appearance [9] and using their skeletons [10]. Polana and Nelson [11] used local motion statistics computed for image grid cells to classify various types of activities. An approach using the temporal templates and motion history images (MHI) for action representation and classification was suggested by Davis and Bobick in Ref. [12]. Cutler and Davis [13] describe a system for real-time moving object classification based on periodicity analysis. Describing the whole spectrum of papers published in this field is beyond the scope of this paper, and we refer the reader to the surveys [14–16].

Related to our approach are the works of Yacoob and Black [17] and Murase and Sakai [18]. In Ref. [17], human activities are recognized using a parameterized representation of measurements collected during one motion period. The measurements are eight motion parameters tracked for five body parts (arm, torso, thigh, calf and foot).

Murase and Sakai [18] describe a method for moving object recognition for gait analysis and lip reading. Their method is based on parametric eigenspace representation of the object silhouettes. Moving object segmentation is performed by using simple background subtraction. The eigenbasis is computed for one class of objects. Murase and Sakai used an optimization technique for finding the shift and scale parameters which is not trivial to implement.

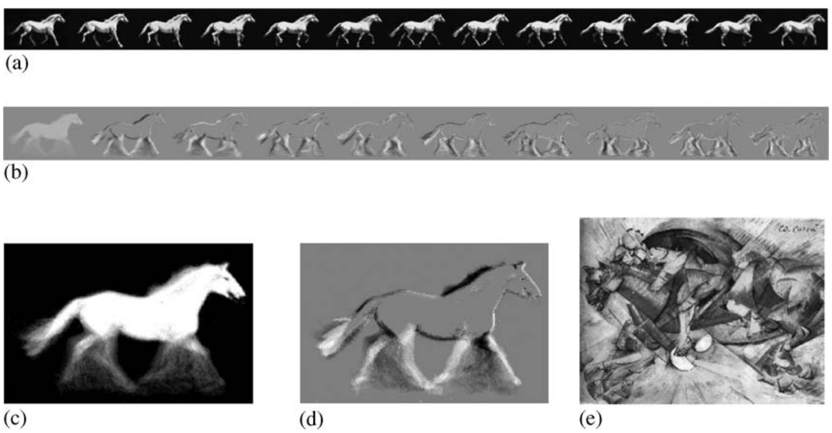

Fig. 2. (a) Running horse video sequence, (b) first 10 eigenshapes, (c,d) first and second eigenshapes enlarged, (e) 'The Red Horseman', 1914, by Carlo Carra, Civico Museo d'Arte Contemporanea, Milan.

In this paper we concentrate on the analysis of deformations of moving non-rigid bodies. We thereby attempt to extract characteristics that allow us to distinguish between different types of motions and different classes of objects.

Essentially we build our method based on approaches suggested in Refs. [18,17] by revisiting and improving some of the components. Unlike previous related methods, we first introduce an advanced variational model for accurate segmentation and tracking. The proposed segmentation and tracking method allows extraction of moving object contours with high accuracy and, yet, is computationally efficient enough to be implemented in a real time system. The algorithm analyzing high-precision object contours achieves better classification results and can be easier adopted to various object classes than the methods based on control points tracking [17] or background subtraction [18].

We then compute eigenbasis for several classes of objects (e.g. humans and animals). Moreover, at this first stage we select the right class by comparing the distance to the right feature space in each of the eigenbases. Next, we create a principle basis not only for static object view (single frames), but also for the whole period subsequences. For example, we create a principle basis for one human motion step and then project every period onto this basis. Unlike [18], we are not trying to build a separate basis for every motion type, but instead, build a common basis for the whole human motion class. Different types of human motions are then detected by looking at the projections of motion samples onto this basis. Apparently, this type of analysis can only be possible due to the high quality of segmentation and tracking results.

We propose an explicit temporal alignment process for finding the shift and scale. Finally, we use a simple classification approach to demonstrate the performances of our framework and test it on various classes of moving objects.

3. Our approach

Our basic assumption is that for any given class of moving objects, like humans, dogs, cats, and birds, the apparent object view in every phase of its motion can be encoded as a combination of several basic body views or configurations. Assuming that a living creature exhibits a pseudo-periodic motion, one motion period can be used as a comparable information unit. The basic views are then extracted from a large training set. An observed sequence of object views collected from one motion period is projected onto this basis. This forms a parameterized representation of the object's motion that can be used for classification.

Unlike [17] we do not assume an initial segmentation of the body into parts and do not explicitly measure the motion parameters. Instead, we work with the changing apparent view of deformable objects and use the parameterization induced by their form variability.

In what follows we describe the main steps of the process that include,

- Segmentation and tracking of the moving object that yield an accurate external object boundary in every frame.
- Periodicity analysis, in which we estimate the frequency of the pseudo-periodic motion and split the video sequence into single-period intervals.
- Frame sequence alignment that brings the single-period sequences to a standardized form by compensating for temporal shift, speed variations, different object sizes and imaging conditions.
- Parameterization by building an eigenshape basis from a training set of possible object views and projecting the apparent view of a moving body onto this basis.

3.1. Segmentation and tracking

Our approach is based on the analysis of deformations of the moving body. Therefore the accuracy of the segmentation and tracking algorithm in finding the target outline is crucial for the quality of the final result. This rules out a number of available or easy-to-build tracking systems that provide only a center of mass or a bounding box around the target and calls for more precise and usually more sophisticated solutions.

Therefore, we decided to use the geodesic active contour approach [19] and specifically the 'fast geodesic active contour' method described in Ref. [2]. The segmentation problem is expressed as a geometric energy minimization. We search for a curve C that minimizes the functional

$$S[\mathscr{C}] = \int_0^{L(\mathscr{C})} g(\mathscr{C})\,\mathrm{d}s,$$

where ds is the Euclidean arclength, $L(\mathscr{C})$ is the total Euclidean length of the curve, and g is a positive edge indicator function in a 3-D hybrid spatial-temporal space that depends on the pair of consecutive frames $I^{t-1}(x, y)$ and $I^t(x, y)$, where $I^t(x, y) : \Omega \subset R^2 \times [0, T] \to R^+$. It gets small values along the spatial-temporal edges, i.e. moving object boundaries, and higher values elsewhere.

In addition to the scheme described in Ref. [2], we also use the background information whenever a static background assumption is valid and a background image $B(x, y)$ is available. In the active contours framework this can be achieved either by modifying the g function to reflect the edges in the difference image $D(x, y) = |B(x, y) - I^t(x, y)|$, or by introducing additional area integration terms to the functional $S(\mathscr{C})$:

$$S[\mathscr{C}] = \int_0^{L(\mathscr{C})} g(\mathscr{C})\,\mathrm{d}s + \lambda_1 \int_{\Omega_C} |D(x, y) - c_1|^2\,\mathrm{d}a + \lambda_2 \int_{\Omega \setminus \Omega_C} |D(x, y) - c_2|^2\,\mathrm{d}a,$$

Fig. 3. Non-rigid moving object segmentation and tracking.

where Ω_C is the area inside the contour C, λ_1 and λ_2 are fixed parameters and c_1, c_2 are given by

$$c_1 = \text{average}_{\Omega_C}[D(x, y)],$$
$$c_2 = \text{average}_{\Omega \setminus \Omega_C}[D(x, y)].$$

The latter approach is inspired by the 'active contours without edges' model proposed by Chan and Vese [20]. It forces the curve $\mathscr{C}$ to close on a region whose interior and exterior have approximately uniform $D(x, y)$ values. A different approach to utilize the region information by coupling between the motion estimation and the tracking problem was suggested by Paragios and Deriche in Ref. [21].

Fig. 3 shows some results of moving object segmentation and tracking using the proposed method.

Contours can be represented in various ways. Here, in order to have a unified coordinate system and be able to apply a simple algebraic tool, we use the implicit representation of a simple closed curve as its binary image. That is, the contour is given by an image for which the exterior of the contour is black while its interior is white.

3.2. Periodicity analysis

Here we assume that the majority of non-rigid moving objects are self-propelled alive creatures, whose motion is almost periodic. Thus, one motion period, like a step of a walking man or a rabbit hop, can be used as a natural unit of motion. Extracted motion characteristics can by normalized by the period size.

The problem of detection and characterization of periodic activities was addressed by several research groups and the prevailing technique for periodicity detection and measurements is the analysis of the changing 1-D intensity signals along spatio-temporal curves associated with a moving object or the curvature analysis of feature point trajectories [22–25]. Here we address the problem using global characteristics of motion such as moving object contour deformations and the trajectory of the center of mass.

By running frequency analysis e.g. Refs. [13,26] on such 1-D contour metrics as the contour area, velocity of the center of mass, principal axes orientation, etc. we can detect the basic period of the motion. Figs. 4 and 5 present global motion characteristics derived from segmented moving objects in two sequences. One can clearly observe the common dominant frequency in all three graphs.

The period can also be estimated in a straightforward manner by looking for the frame where the external object contour best matches the object contour in the current frame. Fig. 6 shows the deformations of a walking man contour during one motion period (step). Samples from two different steps are presented and each vertical pair of frames is phase synchronized. One can clearly see the similarity between the corresponding contours. An automated contour matching can be performed in a number of ways, e.g. by comparing contour signatures or by looking at the correlation between the object silhouettes in different frames as we chose to do in our work. Fig. 7 shows four graphs of inter-frame silhouette correlation values measured for four different starting frames taken within one motion period. It is clearly visible that all four graphs nearly coincide and the local maxima peaks are approximately evenly spaced. The period, therefore, was estimated as the average distance between the neighboring peaks.

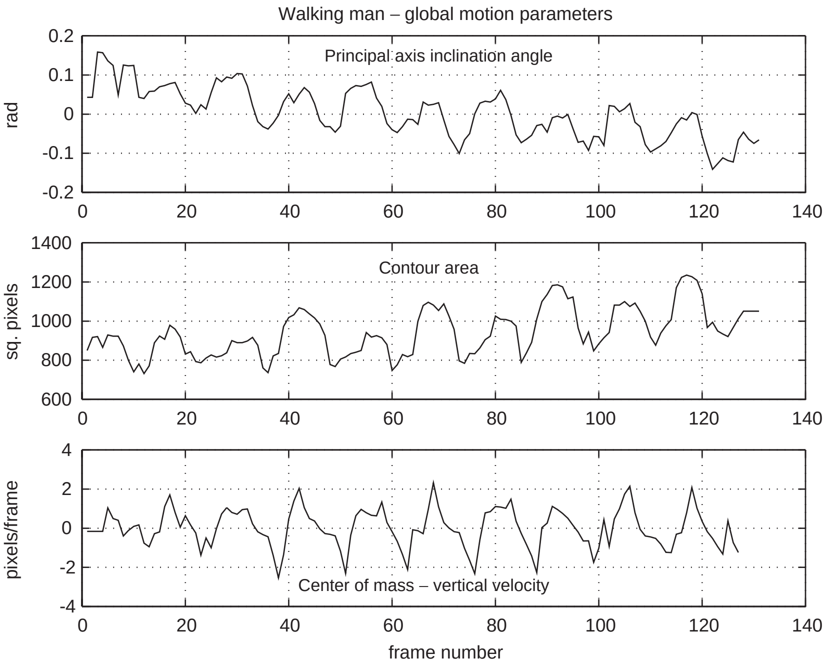

Fig. 4. Global motion characteristics measured for walking man sequence.

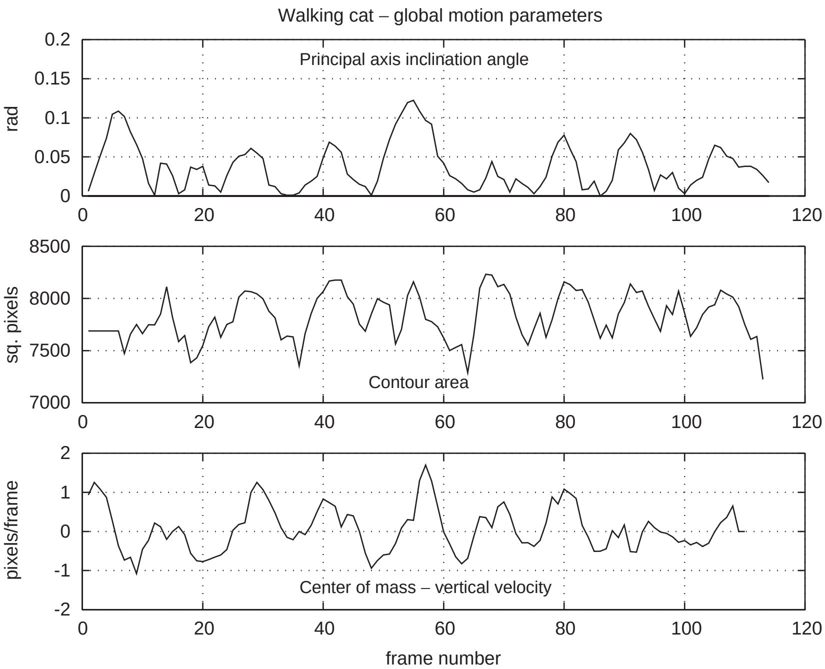

Fig. 5. Global motion characteristics measured for walking cat sequence.

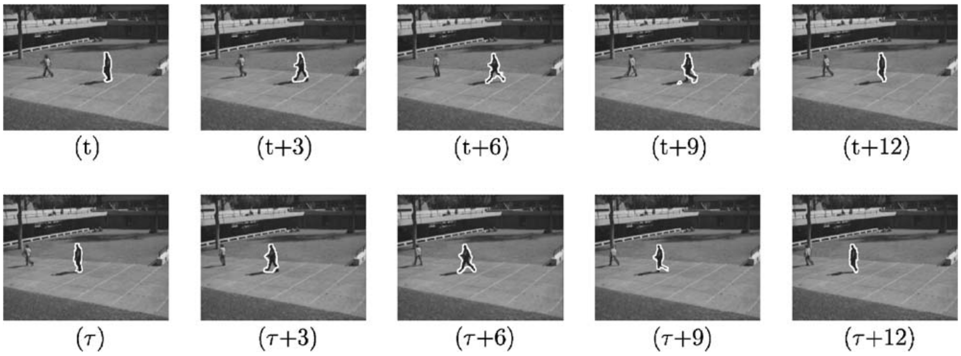

Fig. 6. Deformations of a walking man contour during one motion period (step). Two steps synchronized in phase are shown. One can see the similarity between contours in corresponding phases.

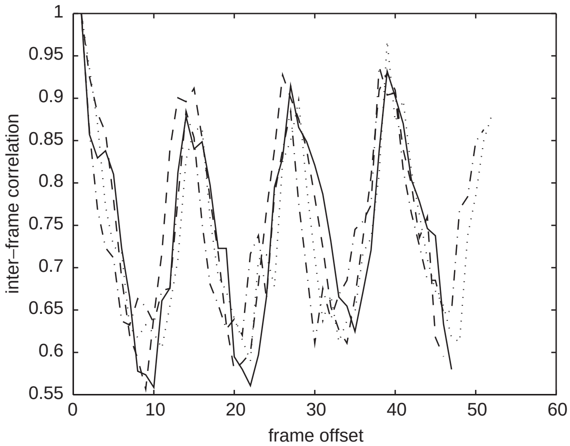

Fig. 7. Inter-frame correlation between object silhouettes. Four graphs show the correlation measured for four initial frames.

3.3. Frame sequence alignment

One of the most desirable features of any classification system is the invariance to a set of possible input transformations. As the input in our case is not a static image, but a sequence of images, the system should be robust to both spatial and temporal variations. In the section below we describe the methods for spatial and temporal alignment of input video sequence bringing it to a canonical, comparable form allowing later classification.

3.3.1. Spatial alignment

We want to compare every object to a single pattern regardless of its size, viewing distance and other imaging conditions. This is achieved by cropping a square bounding box around the center of mass of the tracked target silhouette and re-scaling it to a predefined size (see Fig. 8). Note that this is not a scale invariant transformation unlike other reported normalization techniques e.g. Ref. [27].

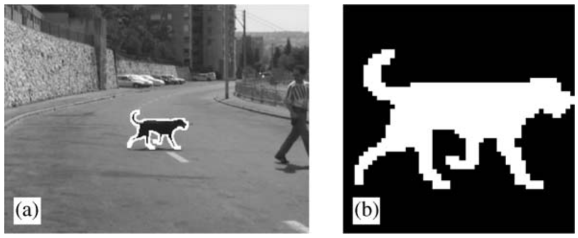

Fig. 8. Scale alignment. A minimal square bounding box around the center of the segmented object silhouette (a) is cropped and re-scaled to form a 50×50 binary image (b).

One way to have orientation invariance is to keep a collection of motion samples for a wide range of possible motion directions and then look for the best match. This approach was used by Yacoob and Black in Ref. [17] to distinguish between different walking directions. Although here we experiment only with motions nearly parallel to the image plane, the system proved to be robust to small variations in orientation. Since we do not want to keep models for both left-to-right and right-to-left motion directions, the right-to-left moving sequences are converted to left-to-right by horizontal mirror flip.

3.3.2. Temporal alignment

A good estimate of the motion period allows us to compensate for motion speed variations by re-sampling each period subsequence to a predefined duration. This can be done by interpolation between the binary silhouette images themselves or between their parameterized representation as explained below. Fig. 9 presents an original and re-sampled one-period subsequence after scaling from 11 to 10 frames.

Temporal shift is another issue that has to be addressed in order to align the phase of the observed one-cycle sample and the models stored in the training base. In Ref. [17] it was done by solving a minimization problem of finding the

Fig. 9. Temporal alignment. Top: original 11 frames of one period subsequence. Bottom: re-sampled 10 frames sequence.

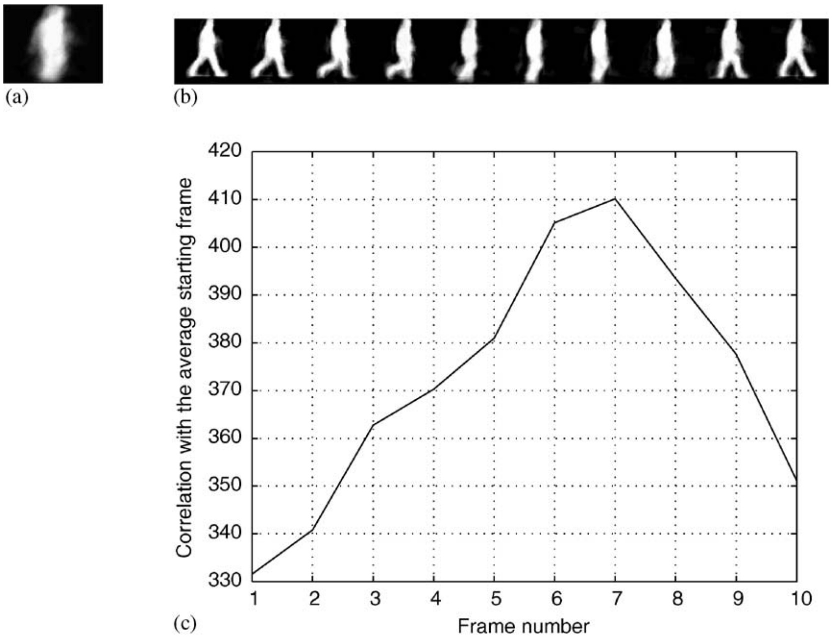

Fig. 10. Temporal shift alignment: (a) average starting frame of all the training set sequences, (b) temporally shifted single-cycle test sequence, (c) correlation between the reference starting frame and the test sequence frames.

optimal parameters of temporal scaling and time shift transformations so that the observed sequence is best matched to the training samples. Polana and Nelson [11] handled this problem by matching the test one-period subsequence to reference template at all possible temporal translations. Murase and Sakai [18] used an optimization technique for finding the shift and scale parameters. Here we propose an explicit temporal alignment process.

Let us assume that in the training set all the sequences are accurately aligned. Then we find the temporal shift of a test sequence by looking for the starting frame that best matches the generalized (averaged) starting frame of the training samples, as they all look alike. Fig. 10 shows (a) the reference starting frame taken as an average over the temporally aligned training set, (b) a re-sampled single-period test sequence and, (c) the correlation between the reference starting frame and the test sequence frames. The maximal correlation is achieved at the seventh frame, therefore the test sequence is aligned by cyclically shifting it 7 frames to the left.

3.4. Parameterization

In order to reduce the dimensionality of the problem we first project the object image in every frame onto a low-dimensional base. The base is chosen to represent all possible appearances of objects that belong to a certain class, like humans, four-leg animals, etc.

Let n be number of frames in the training base of a certain class of objects and M be a training samples matrix, where each column corresponds to a spatially aligned image of a moving object written as a binary vector. In our experiments we use 50×50 normalized images, therefore, M is a $2500 \times n$ matrix. The correlation matrix MM^{T} is decomposed using Singular Value Decomposition as $MM^{\mathrm{T}} = U\Sigma V^{\mathrm{T}}$, where U is an orthogonal matrix of principal directions and the Σ is

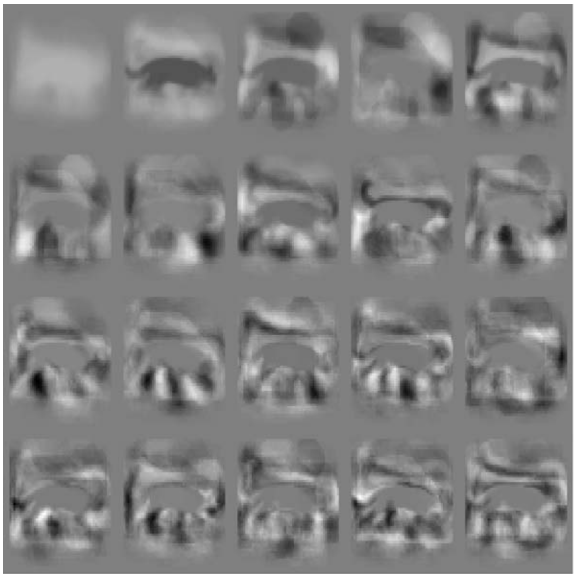

Fig. 11. The principal basis for the ‘dogs and cats’ training set formed of 20 first principal component vectors.

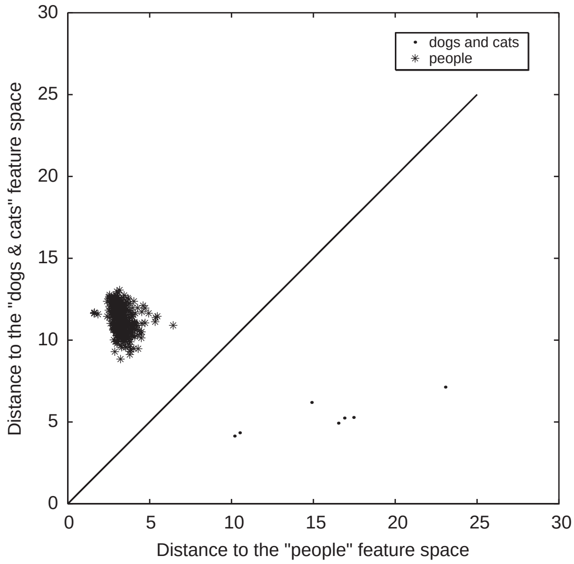

Fig. 12. Distances to the ‘people’ and ‘dogs and cats’ feature spaces from more than 1000 various images of people, dogs and cats.

a diagonal matrix of singular values. In practice, the decomposition is performed on $M^{\mathrm{T}}M$, which is computationally more efficient [28]. The principal basis $\{U_i, i=1,\ldots,k\}$ for the training set is then taken as k columns of U corresponding to the largest singular values in Σ. Fig. 11 presents a principal basis for the training set formed of 800 sample images collected from more than 60 sequences showing dogs and cats in motion. The basis is built by taking the $k=20$ first principal component vectors.

We build such representative bases for every class of objects. Then, we can distinguish between various object classes by finding the basis that best represents a given object image in the distance to the feature space (DTFS) sense. Fig. 12 shows the distances from more than 1000 various images of people, dogs and cats to the feature space of people and to that of dogs and cats. In all cases, images of people were closer to the people feature space than to the animals' feature space and vise a versa. This allows us to distinguish between these two classes. A similar approach was used in Ref. [29] for the detection of pedestrians in traffic scenes.

If the object class is known (e.g. we know that the object is a dog), we can parameterize the moving object silhouette image I in every frame by projecting it onto the class basis. Let B be the basis matrix formed from the basis vectors $\{U_i, i=1,\ldots,k\}$. Then, the parameterized representation of the object image I is given by the vector $\overline{p}$ of length k as $\overline{p}=B^{\mathrm{T}}\overline{v_I}$, where $\overline{v_I}$ is the image I written as a vector.

The idea of using a parameterized representation in motion-based recognition context is certainly not a new one. To name a few examples we mention again the works of Yacoob and Black [17], and of Murase and Sakai [18]. Cootes et al. [30] used similar technique for describing feature point locations by a reduced parameter set. Baumberg and Hogg [31] used PCA to describe a set of admissible B-spline models for deformable object tracking. Chomat and Crowley [32] used PCA-based spatio-temporal filter for human motion recognition.

Fig. 13 shows several normalized moving object images from the original sequence and their reconstruction from a parameterized representation by back-projection to the image space. The numbers below are the norms of differences between the original and the back-projected images. These norms can be used as the DTFS estimation.

Now, we can use these parameterized representations to distinguish between different types of motion. The reference base for the activity recognition consists of temporally aligned one-period subsequences. The moving object silhouette in every frame of these subsequences is represented by its projection to the principal basis. More formally, let $\{I_f : f=1,\ldots,T\}$ be a one-period, temporally aligned set of normalized object images, and $\overline{p_f}$, $f=1,\ldots,T$ a projection of the image I_f onto the principal basis B of size k. Then, the vector P of length kT formed by concatenation of all the vectors $\overline{p_f}$, $f=1,\ldots,T$, represent a one-period subsequence. By choosing a basis of size $k=20$ and the normalized duration of one-period subsequence to be $T=10$ frames, every single-period subsequence is represented by a feature point in a 200-dimensional feature space.

In the following experiment we processed a number of sequences of dogs and cats in various types of locomotion. From these sequences we extracted 33 samples of walking dogs, 9 samples of running dogs, 9 samples with galloping dogs and 14 samples of walking cats. Let $S_{200\times m}$ be

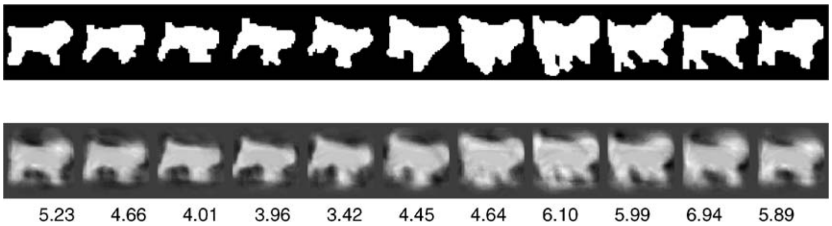

Fig. 13. Image sequence parameterization. Top: 11 normalized target images of the original sequence. Bottom: the same images after the parameterization using the principal basis and back-projecting to the image basis. The numbers are the norms of the differences between the original and the back-projected images.

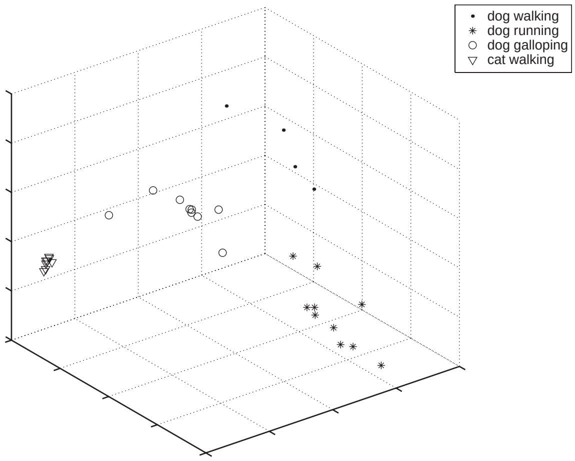

Fig. 14. Feature points extracted from the sequences with walking, running and galloping dogs and walking cats and projected to the 3D space for visualization.

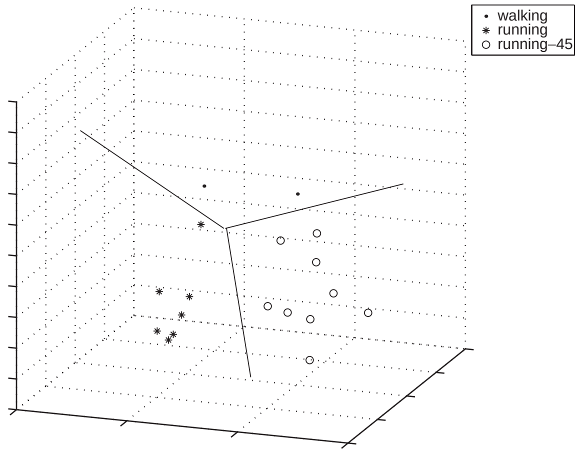

Fig. 15. Feature points extracted from the sequences showing people walking and running parallel to the image plane and at 45° angle to the camera. Feature points are projected to the 3-D space for visualization.

the matrix of projected single-period subsequences, where m is the number of samples and the SVD of the correlation matrix is given by $SS^{\mathrm{T}} = U^S \Sigma^S V^S$. In Fig. 14 we depict the resulting feature points projected for visualization to the 3-D space using the three first principal directions $\{U_i^S : i = 1, \ldots, 3\}$, taken as the column vectors of U^S corresponding to the three largest eigen-values in Σ^S. One can easily observe four separable clusters corresponding to the four groups.

Another experiment was done over the 'people' class of images. Fig. 15 presents feature points corresponding to several sequences showing people walking and running parallel to the image plane and running at oblique angle to the camera. Again, all three groups lie in separable clusters.

The classification can be performed, for example, using the k-nearest-neighbor algorithm. We conducted the 'leave one out' test for the dogs set above, classifying every sample by taking them out from the training set one at a time. The three-nearest-neighbors strategy resulted in 100% success rate.

Another option is to further reduce the dimensionality of the feature space by projecting the 200-dimensional feature vectors in S to some principal basis. This can be done using a basis of any size, exactly as we did in 3-D for visualization. Fig. 16 presents learning curves for both animals and people data sets. The curves show how the classification error rate achieved by one-nearest-neighbor classifier changes as a function of principal basis size. The curves are obtained by averaging over a hundred iterations when the data set is randomly split into training and testing sets. The principal basis is built on the training set and the testing is performed using the 'leave one out' strategy on the testing set.

4. Concluding remarks

We presented a new framework for motion-based classification of moving non-rigid objects. The technique is based on the analysis of changing appearance of moving objects and is heavily relying on high accuracy results of segmentation and tracking by using the fast geodesic contour

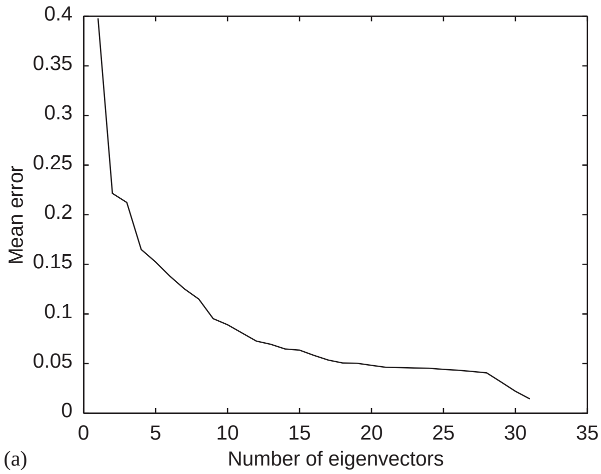

Fig. 16. Learning curves for (a) dogs and cats data set and (b) people data set. One-nearest-neighbor classification error rate as a function of the number of eigenvectors in the projection basis.

approach. The periodicity analysis is then performed based on the global properties of the extracted moving object contours, followed by video sequence spatial and temporal normalization. Normalized one-period subsequences are parameterized by projection onto a principal basis extracted from a training set of images for a given class of objects. A number of experiments show the ability of the system to analyze motions of humans and animals, to distinguish between these two classes based on object appearance, and to classify various type of activities within a class, such as walking, running, galloping. The 'dogs and cats' experiment demonstrate the ability of the system to discriminate between these two very similar by appearance classes by analyzing their locomotion.

Acknowledgements

We thank Alfred M. Bruckstein for sharing with us his observation of the connection between the Futurism art movement and the periodic motion presentation using eigenshapes.

References

[1] J.R. Beniger, The Arts and New Media site. In www.usc.edu/schools/annenberg/asc/projects/comm544/, University of South California, Annenberg School for Communication.

[2] R. Goldenberg, R. Kimmel, E. Rivlin, M. Rudzsky, Fast geodesic active contours, IEEE Trans. Image Process. 10 (10) (2001) 1467–1475.

[3] K. Gould, M. Shah, The trajectory primal sketch: a multi-scale scheme for representing motion characteristics, in: Proceedings of the Conference on Computer Vision and Pattern Recognition, San Diego, CA, 1989, pp. 79–85.

[4] K. Gould, K. Rangarajan, M. Shah, Detection and representation of events in motion trajectories, in: Gonzalez, Mahdavieh (Eds.), Advances in Image Processing and Analysis, SPIE Optical Engineering Press, June 1992, (Chapter 14).

[5] S.A. Engel, J.M. Rubin, Detecting visual motion boundaries, in: Proceedings of the Workshop on Motion: Representation and Analysis, Charleston, SC, May 1986, pp. 107–111.

[6] J.M. Siskind, Q. Morris, A maximum-likelihood approach to visual event classification, in: Proceedings of the Fourth European Conference on Computer Vision, Cambridge, UK, April 1996, pp. 347–360.

[7] D. Koller, K. Daniilidis, H.-H. Nagel, Model-based object tracking in monocular image sequences of road traffic scenes, Int. J. Comput. Vision 10 (1993) 257–281.

[8] N.J. Cutaia, J.A. O'Sullivan, Automatic target recognition using kinematic priors, in: Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, FL, December 1994, pp. 3303–3307.

[9] A. Lipton, H. Fujiyoshi, R. Patil, Moving target classification and tracking from real-time video, in: Proceedings of the IEEE Image Understanding Workshop, 1998, pp. 129–136.

[10] H. Fujiyoshi, A. Lipton, Real-time human motion analysis by image skeletonization, in: Proceedings of the Workshop on Application of Computer Vision, October 1998.

[11] R. Polana, R.C. Nelson, Detection and recognition of periodic, nonrigid motion, Int. J. Comput. Vision 23 (3) (1997) 261–282.

[12] A. Bobick, J. Davis, The representation and recognition of action using temporal templates, IEEE Trans. Pattern Anal. Mach. Intell. 23 (3) (2001) 257–267.

[13] R. Cutler, L. Davis, Robust real-time periodic motion detection, analysis, and applications, Pattern Anal. Mach. Intell. 22 (8) (2000) 781–796.

[14] C. Cedras, M. Shah, Motion-based recognition: a survey, Image Vision Comput. 13 (2) (1995) 129–155.

[15] D.M. Gavrila, The visual analysis of human movement: a survey, Comput. Vision Image Understanding 73 (1) (1999) 82–98.
[16] D.M. Moeslund, E. Granum, A survey of computer vision-based human motion capture, Comput. Vision Image Understanding 81 (3) (2001) 231–268.
[17] Y. Yacoob, M.J. Black, Parameterized modeling and recognition of activities, Comput. Vision Image Understanding 73 (2) (1999) 232–247.
[18] H. Murase, R. Sakai, Moving object recognition in eigenspace representation: gait analysis and lip reading, Pattern Recogn. Lett. 17 (1996) 155–162.
[19] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, Int. J. Comput. Vision 22 (1) (1997) 61–79.
[20] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Process. 10 (2) (2001) 266–277.
[21] N. Paragios, R. Deriche, Geodesic active regions for motion estimation and tracking, in: Proceedings of the Seventh International Conference on Computer Vision, Kerkyra, Greece, 1999, pp. 688–694.
[22] R. Polana, R.C. Nelson, Detecting activities, J. Visual Commun. Image Representat. 5 (1994) 172–180.
[23] F. Liu, R.W. Picard, Finding periodicity in space and time, in: Proceedings of the Sixth International Conference on Computer Vision, Bombay, India, 1998, pp. 376–383.
[24] S.M. Seitz, C.R. Dyer, View invariant analysis of cyclic motion, Int. J. Comput. Vision 25 (3) (1997) 231–251.
[25] P.S. Tsai, M. Shah, K. Keiter, T. Kasparis, Cyclic motion detection for motion based recognition, Pattern Recognition 27 (12) (1994) 1591–1603.
[26] C. Rao, A. Yilmaz, M. Shah, View-invariant representation and recognition of actions, Int. J. Comput. Vision 50 (2) (2002) 203–226.
[27] C. Stauffer, K. Tieu, L. Lee, Robust automated planar normalization of tracking data, in: Proceedings of Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2003.
[28] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neuro Sci. 3 (1) (1991) 71–86.
[29] U. Franke, D. Gavrila, S. Gorzig, F. Lindner, F. Paetzold, C. Wohler, Autonomous driving goes downtown, IEEE Intell. System 13 (6) (1998) 40–48.
[30] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active shape models: their training and application, Comput. Vision Image Understanding 61 (1) (1995) 38–59.
[31] A. Baumberg, D. Hogg, An efficient method for contour tracking using active shape models, in: Proceedings of the IEEE Workshop on Motion of Non-Rigid and Articulated Objects, Austin, 1994, pp. 194–199.
[32] O. Chomat, J. Crowley, Recognizing motion using local appearance, in: Proceedings of the International Symposium on Intelligent Robotic Systems, University of Edinburgh, 1998.

About the Author—ROMAN GOLDENBERG received the B.A. (Summa cum Laude) and Ph.D. degrees in Computer Science, from the Technion—Israel Institute of Technology, Haifa, in 1995 and 2003, respectively.

From 1994 to 1996, 1999 he was with IBM Research Lab in Haifa. Currently he is a research fellow at the Technion R&D Foundation and the Samuel Neeman Institute for Advanced Studies in Science and Technology.

His research interests include video analysis, tracking, motion-based recognition, PDE methods for image processing, and medical imaging.

About the Author—RON KIMMEL received his B.Sc. (with honors) in computer engineering in 1986, the M.S. degree in 1993 in electrical engineering, and the D.Sc. degree in 1995 from the Technion—Israel Institute of Technology. During the years 1986–1991 he served as an R&D officer in the Israeli Air Force.

During the years 1995–1998 he has been a postdoctoral fellow at Lawrence Berkeley National Laboratory, and the Mathematics Department, University of California, Berkeley. Since 1998, he has been a faculty member of the Computer Science Department at the Technion, Israel.

His research interests are in computational methods and their applications including topics in: Differential geometry, numerical analysis, non-linear image processing, geometric methods in computer vision, and numerical geometry methods in computer aided design, robotic navigation, and computer graphics.

Dr. Kimmel was awarded the Hershel Rich Technion innovation award, the Henry Taub Prize for excellence in research, Alon Fellowship, the HTI Post-doctoral Fellowship, and the Wolf, Gutwirth, Ollendorff, and Jury fellowships. He has been a consultant of HP research Lab in image processing and analysis during the years 1998–2000, and to Net2Wireless/Jigami research group during 2000–2001.

About the Author—EHUD RIVLIN received the B.Sc. and M.Sc. degrees in computer science and the M.B.A. degree from the Hebrew University in Jerusalem, and the Ph.D. from the University of Maryland.

Currently he is an Associate Professor in the Computer Science Department at the Technion, Israel Institute of Technology. His current research interests are in machine vision and robot navigation.

About the Author—MICHAEL RUDZSKY received his Ph.D. in physics and mathematics from the Institute of Space Research, Moscow in 1980. He worked in the Scientific and Industrial Association for Space Research in Baku, Azerbajan till 1990.

Since 1991 he was a research fellow in the Physics Department and since 1995 at the Computer Science Department of the Technion, Haifa, Israel. His current research interests include computer vision, pattern recognition and compression of images.