Classification of upper limb motion trajectories using shape features

Huiyu Zhou, Huosheng Hu, Senior Member, IEEE, Honghai Liu, Senior Member, IEEE, and Jinshan Tang, Senior Member, IEEE

Abstract—Understanding and interpreting human motion is a very active research area nowadays due to its importance in sports sciences, health care and video surveillance. However, classification of human motion patterns is still a challenging topic due to the variations in kinetics and kinematics of human movements. This paper presents a novel algorithm for automatic classification of motion trajectories of human upper limbs. The proposed scheme starts from transforming 3-D positions and rotations of the shoulder/elbow/wrist joints into 2-D trajectories. Discriminative features of these 2-D trajectories are then extracted using a probabilistic shape context method. Afterwards, these features are classified using a k-means clustering algorithm. Experimental results demonstrate the superiority of the proposed method against the state of the art techniques.

Index Terms—Motion trajectory, classification, health care, shape contexts, expectation-maximisation.

I. INTRODUCTION

In the last decade, ageing population in developed countries have rapidly increased. It is necessary to provide people immediate access to high quality, affordable and accessible health care services any place and anytime without increasing costs. This challenge inspires fully functional electronic designs that ensure accurate, automatic and efficient delivery of health care services. Solutions like telemedicine can help reduce the duration of hospital stays and the need for frequent visits to a hospital. One of the active research areas in telemedicine is to provide remote monitoring services, where human behaviours are fully analysed so that appropriate medical treatments can be delivered.

Understanding and interpreting human motion, based on video sequences, is an active research area with popular applications in sports sciences, medicine, biomechanics and surveillance (e.g. [8], [15], [36]). Human motion analysis is mainly driven by the developments in computer graphics, computer vision and biomechanics, e.g. [15], [16]. In recent years, studies on kinematics and dynamics of human upper limbs have become overwhelming in medical science (e.g. [38]), as upper extremity dysfunction strongly influences the capacity of a person to perform activities of daily life (ADL) such as self-care, dressing, eating and object manipulations [10].

It has been reported that appropriate aids and treatments carefully prescribed can enhance the recovery of arm mobility among the patients [6]. Therefore, a precise, objective and sensitive quantification of dysfunction of upper limbs is a prerequisite before a care plan is generated [34]. This quantification also allows clinicians to better understand the nature of the disease. Current strategies for assessing functional motion abilities of upper limbs are limited to subjective evaluations, conducted by clinicians working in hospital, using the criteria such as dexterity and speed of single-hand movements, dexterity and speed of both hands [33].

These test criteria, in spite of their simplicity and efficiency, tend to be lack of robustness. For example, the impairment index of Parkinson’s disease can vary by up to 40% among different observers [44]. As a promising solution towards this problem, camera based intelligent decision systems have been developed to recognise the movements of human upper limbs (e.g. [38]). For example, Petuskey et al. [30] established a normative pediatric database of 3-D kinematic values during selected ADLs, where children with limited upper extremity functions and ADL performance can be thoroughly studied.

Among the established work for human motion analysis, shape based classification approaches recently draw much attention of the research community, e.g. [39], [45]. Dynamic time warping (DTW) is a template-based dynamic programming algorithm for shape matching. This technique is of conceptual simplicity and stability [25]. However, it is computationally demanding and requires string alignments in the starting point. Shape context is a feature based scheme that utilises the histogram of the relative polar coordinates of all other points [3]. Shape context also needs high computational efforts but can be used to describe the similarity of two deformed shapes. Other commonly used shape descriptors are Minimum Description Length, Fourier analysis, curvature scale space, Zernike moment, structural information and patch and skeleton based computation, e.g. [24].

This paper demonstrates how motion trajectories of human upper limbs are automatically classified, based on the measures of shape similarity calculated beforehand. To our knowledge, most of the existing feature based classifiers rarely take into account the association between two images in the stage of feature extraction. In many cases, this practice very likely experiences large uncertainty in the classification stage, as the extracted features may not fully describe the images and
inherent characteristics.

Our work is to explore the perceptual difference using an iterative parsing process, where shape context is used to calculate shape distance after two shapes have been somehow registered. Shape context is used here due to its robustness performance [3]. The main contribution of our approach is the extension of shape context in combination with a probabilistic scheme, which ensures two shapes to be well registered. This “similarity maximisation” approach, in spite of its computational efforts, is used to improve the discrimination capability of the classical shape context based methods in different motion circumstances. The applications of the proposed approach include human action analysis and understanding in video surveillance.

The entire paper is organised as follows. In the next section, related work with the applications in biomedical engineering is summarised. In Section III, a classical shape context approach and its properties will be introduced. A new shape matching strategy based on the expectation-maximisation algorithm is reported in Section IV. Afterwards, the evaluation of the proposed strategy is performed in Section V. Finally, conclusions and future work will be given in Section VI.

II. RELATED WORK

In this section, the state of the art techniques developed for classification and recognition of the upper limb movements are summarised. Particularly, the focus is on the applications of these classical approaches in biomedical fields. Firstly, the feature- and evidence-based approaches for motion analysis of the upper extremity are addressed. Afterwards, the established classification approaches used in the mobility measurements will be introduced.

A. Applications of feature based motion analysis

Lee et al. [18] presented an upper-limb-movement classification system for measuring the impairment degrees of cerebral palsy children using momentum analysis. Visual tracking tests were used to obtain a quantified statistical description of the involuntary movements of the arm around the elbow joint in a group of patients suffering from athetoid cerebral palsy [26], where frequency components were extracted.

Sanger [31] quantified the variability in arm trajectories of seven children with dyskinetic cerebral palsy through repeated outward reaching movements. It was observed that children with dyskinetic cerebral palsy had a significantly reduced signal-to-noise ratio compared with healthy children at similar ages. Tsao and Mirbagheri [37] studied upper-extremity movements in individuals with neurological disorders such as stroke and spinal cord injury (SCI) by measuring movement trajectories and their derivatives. Zhang and Chaffin [47] presented a study to quantitatively assess the effects of speed variations on the joint kinematics during multisegment-seated reaching movements.

Gordon et al. [12] performed quantitative measures so as to assess spasticity and dystonia. Spasticity was measured as the slope of the force-velocity relationship from a test when one measures the forces required to passively extend the elbow at different velocities. Duric et al. [11] proposed a method for tracking and recognition of the lower arms and hand movements from video sequences using a linguistic approach.

B. Classification techniques for mobility measurement

In [41], an approach for automatically segmenting sequences of natural activities into atomic sections and finally clustering them was presented. The continuous observations were transformed to a discrete symbol sequence. Integration of spatial and temporal contexts that require a discrete sequence of symbol vectors have been used for action recognition, e.g. [32].

A view-independent approach was presented to the recognition of human gestures in a low resolution sequence from multiple calibrated cameras [5]. Hu moments were used in [4] for the matching of temporal templates. A visual hull was reconstructed by computing the polyhedral approximations of the silhouettes [9]. Weinland et al. [43] proposed a new framework which modeled actions using three dimensional occupancy grids as visual hulls, built from multiple viewpoints in an exemplar-based Hidden Markov Model (HMM). Hahn et al. [13] introduced a novel trajectory classification approach, depending on the Levenshtein Distance on Trajectories (LDT) as a measure of the similarity between two motion trajectories.

Lu and Ferrier [20] presented a two-threshold segmentation algorithm to automatically decompose a complex motion into a sequence of simple linear dynamic models. A compact motion representation was obtained for each segment using the parameters of a damped harmonic dynamic model [1]. Kim and Cipolla [17] proposed an action recognition method by extending Canonical correlation analysis (CCA) of two sets of vectors into two video volume tensors. Tensor based approaches have also been integrated in e.g. [40] for the action recognition purpose. Ali and Shah [2] proposed a set of kinematic features, e.g. divergence, vorticity, symmetric and anti-symmetric flow fields, that are derived from the optical flow for human action recognition in videos.

Recently, shape context has been used as reference points to offer a globally discriminative characterisation for object recognition [3]. Shape context refers to the distribution of a
set of vectors originating from a point to the remaining points on a curve. The similarity/dissimilarity measurements between the two curves, namely “shape distance”, are computed as a sum of matching errors between the corresponding points. The applications of shape context can be found in [19],[42].

III. BACKGROUND

A. Preliminary

Fig. 1 illustrates an example in which motion trajectories of the right wrist with respect to the shoulder position have been obtained for different postures. In this example, the three postures demonstrate certain similar patterns. Mean-square spectrum may be used to separate these different trajectories in case they have different average spectrum values.

In time and frequency domains, other statistical features can also be used to discriminate the measurements, for examples, autoregressive coefficients of order 2 and 6 (AR2 and AR6), integrated absolute value (IAV), mean absolute value (MAV), root mean square (RMS), slope sign change (SSC) and wave-form length (WL) [7],[27]. Table 1 shows individual results after the statistical features have been generated from a set of pull, push and punch patterns.

However, the effectiveness of the statistical features is inconsistent. For example, Fig. 2 shows that the mean-square spectra of the overall motion trajectories look very similar. Moreover, Table II shows that the difference between any two of the calculated statistical features is much smaller than the corresponding items reported in Table I. Especially, if both “pull” and “punch” happen at a similar speed, the statistical features (e.g. IAV, MAV, RMS and SSC) extracted from these motion patterns become less discriminative than they are expected. Therefore, it is necessary to develop a new strategy to maintain the discrimination capability in all the cases.

B. Shape contexts

Classification of two motion trajectories is literally a procedure where the similarity degree of two trajectories is calculated. It is also a correspondence/registration problem: correspondence or registration errors can be used to measure the similarity between them. Shape context is such a special descriptor that it can be used for evaluating the shape similarity. In this paper, for convenience a motion trajectory is called a shape/curve, which is formed as the motion distance of individual joints against a fixed reference.

Consider a point \(a_i\) on the first shape and a point \(b_j\) on the second shape, where \((i,j)\) indicate the indexes of the points on the two different shapes. The \(\chi\) test statistic can be used to describe the cost of matching these two points:

\[
C_{i,j} = C(a_i, b_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} ,
\]

where \(h_i(k)\) and \(h_j(k)\) denote the \(K\)-bin normalised histograms at \(a_i\) and \(b_j\) respectively. Given the overall corresponding pairs over the two groups, our intention is to minimise the total cost of matching as follows:

\[
H(\gamma) = \sum_i C(a_i, b_{\gamma(i)}) ,
\]

which is subject to the constraint of one-to-one matching (\(\gamma\) is a permutation).

This minimisation procedure is a square assignment (or weighted bipartite matching) problem. To solve this problem one can use the Hungarian method or its variants [29]. As indicated in [3], shape matching must be invariant under various scaling and translation conditions while maintaining robust performance in the presence of certain geometrical distortions, occlusions and outliers.

Suppose that the two shapes can be corresponded subject to a geometrical transformation. Therefore, the relationship in

![Image](https://example.com/image.png)

Fig. 2. Motion trajectories and feature extraction of the right wrist against the right shoulder for different postures: (a) pull, (b) push and (c) punch. Row 1 refers to the trajectories against time, and row 2 indicates that the computed mean-square spectrum of individual motion patterns, which are not distinctive.

![Image](https://example.com/image.png)

Fig. 3. Illustration of shape context. Left: for the deformed model according to “pull” and “push” patterns; right: for the deformed model according to “pull” and “punch” trajectories.
geometry of these two shapes is now investigated. Assume that there is a transformation \( T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \) that can be used to map arbitrary points from one shape to another. An affine model is usually used for the transformation: \( T(x) = Ax + \zeta \), where \( A \) is the transformation matrix and \( \zeta \) is a translational offset vector. A least square solution \( \hat{T}(A, \zeta) \) can be obtained using the following form:

\[
\hat{\zeta} = \frac{1}{n} \sum_{i=1}^{n} (x_i - b_i) \quad \hat{A} = (\phi^+ \phi)^{-1} \phi^+ \xi \quad \text{(3)}
\]

where \( n \) is the number of the extracted feature points, and \( \phi \) and \( \psi \) comprise the homogeneous coordinates of \( \theta \) and \( \psi \) respectively, where

\[
\phi = \begin{bmatrix} a_{11} & a_{12} \\ \vdots & \vdots \\ a_{n1} & a_{n2} \end{bmatrix} \quad \text{(4)}
\]

and \( \phi^+ \) is the pseudoinverse of \( \phi \). The thin plate spline (TPS) model established in [23], due to its stability, can be used to interpret the coordinate transformation. If different locations \((x_i, y_i)\) are noncollinear, the TPS interpolant \( f(x, y) \) is used to minimise the bending energy:

\[
I_f = \int \int_{\mathbb{R}^2} \left( \frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left( \frac{\partial^2 f}{\partial y^2} \right)^2 \, dx \, dy, \quad \text{(5)}
\]

with \( f(x, y) = \sum_{i=1}^{n} w_i U(|| (x, y) - (x_i, y_i) ||) \), where the kernel function \( U(x) \) is defined by \( U(r) = r^2 \log r \) and \( U(0) = 0 \). Let

\[
\sum_{i=1}^{n} w_i = 0, \quad \sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i y_i = 0, \quad \text{(6)}
\]

and \( f(x, y) = v_i \), one can have:

\[
\begin{bmatrix} K & \phi^T \\ \phi & 0 \end{bmatrix} \begin{bmatrix} \phi \\ w/c \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix} \quad \text{(7)}
\]

where \( K_{ij} = U(|| (x_i, y_i) - (x_j, y_j) ||) \), the \( ij \)th row of \( \phi \) is \((1, x_i, y_i)\), \( w \) and \( v \) are column vectors from \( w_i \) and \( v_i \), and \( c \) is the column vector with elements \( c_{11}, c_{12} \) and \( c_p \).

It has been proven that \( I_f \approx w^T K w \) [3]. As pointed out above, shape distance, based on shape context, will be used to measure similarity degrees of two motion trajectories. Generally speaking, shape distance between the two trajectories is calculated using the weighted sum of shape context, image appearance distance and bending energy. The shape distance between shapes \( \phi \) and \( \psi \) is described as:

\[
D_{SC}(\theta, \psi) = \frac{1}{n} \sum_{i=1}^{n} \arg \min_{\theta, \psi} C(a, T(b_i)) + \frac{1}{m} \sum_{b_i \in \psi} \arg \min_{\theta, \psi} C(a, T(b_i)), \quad \text{(8)}
\]

where \( \theta \) and \( \psi \) are the gray level images corresponding to shapes \( \theta \) and \( \psi \) respectively. \( \Delta \) is a differential vector offset and \( G \) is a windowing function (Gaussian). A sum over squared differences within a window around the corresponding points is then used to score the intensity similarity. This score is computed if and only if the thin plate spline transformation \( T \) has been applied to align the two images.

Fig. 3 denotes the calculated shape context of different pairs of motion patterns. It shows that the pair of pull/punch has a significantly different representation from the pair of pull/push. Let us take a look at the shape matching procedure. Fig. 4 illustrates an example of shape matching using the classical shape context. In this example, three motion trajectories (i.e. “pull”, “push” and “punch”) shown in Fig. 2 are used for computing shape similarity. The resultant statistics are tabulated in Table III, which demonstrate that shape context can be used to separate these three patterns.

### IV. PROPOSED APPROACH

Shape context has demonstrated its capability of discriminating similar shapes/motion trajectories. Our main interest is to maintain the performance of shape context based classifiers in different circumstances. For this purpose, the system performance is further investigated after one has applied shape context for similarity measurements. It is evident that in the first graph on the second row of Fig. 4, part of the concave areas on the curves lack necessary correspondences. Ideally, the first two peaks of these two curves are expectedly matched. However, this does not occur in the final outcome. Similar observations are also obtained in the third graph of the second row. This visual discrepancy can be optimally handled if a new shape transformation approach is properly designed. In other words, it is intended to find a better registration strategy.
A. Minimising shape distances

Let \( f(\hat{\theta}_t, \hat{\psi}_t, \tilde{T}_t) \) be a dynamic representation for the shape matching, where \( \hat{\theta}_t \) and \( \hat{\psi}_t \) are the image observations of shapes \( \theta \) and \( \psi \), and \( \tilde{T}_t \) is the estimated shape transformation at time \( t \). If \( \tilde{T}_t \) is optimal, one will be able to match the points of the two shapes using a maximum likelihood estimation, i.e., finding a maximised \( p(\hat{\psi}_t | \tilde{T}_t) \).

Let image points be conditionally independent of each other (as each point indicates an independent event), the following joint probability stands: \( p(\hat{\psi}_t | \tilde{T}_t) = \prod_i p(\hat{\psi}_{t_i} | \tilde{T}_t) \), where \( i \) is the index of one of the image points on the shape. Using Bayes’ rule, one obtains: \( p(\hat{\psi}_{t_i} | \tilde{T}_t) = \sum_j p(\hat{\psi}_{t_i} | \hat{\theta}_{t_j}, \tilde{T}_t) p(\hat{\theta}_{t_j} | \tilde{T}_t) \), where \( j \) is a different image point on the shape. The maximum log-likelihood estimate of \( \tilde{T}_t \) is defined as

\[
L(\tilde{T}_t) \equiv \log p(\hat{\psi}_t | \tilde{T}_t).
\]

Correct correspondence between two sets of image points cannot be known a priori. Therefore, the conditional logarithm of the posterior probability \( Q(\tilde{T}_t | \tilde{T}_{t-1}) \) can be computed over the previous estimates of \( \tilde{T}_t \). Using the Markovian properties, it is expected to maximise the conditional log-likelihood as follows:

\[
Q(\tilde{T}_t | \tilde{T}_{t-1}) = E(\log p(\hat{\psi}_t | \tilde{T}_t) | \tilde{T}_0, ..., \tilde{T}_{t-1}) = \sum_i \sum_j p(\hat{\theta}_{t_j} | \tilde{T}_t) \log p(\hat{\psi}_{t_i} | \hat{\theta}_{t_j}, \tilde{T}_t).
\]

B. Expectation-minimisation algorithm

To estimate the transformation over \( t \), which is the target in this stage, one can use a maximum likelihood estimation (MLE) scheme. Correspondence outliers usually are inevitable and an expectation-maximisation (EM) strategy can be used in order to reach the minimisation of shape distance. The EM algorithm starts from an initial estimate of shape transformation, which is performed using random samples from the image points on the curves. An optimal correspondence is sought over two shapes using an iterative scheme:

1. **E-step**: The a posteriori probability of the incomplete data set, \( p(\hat{\theta}_{t_j} | \hat{\psi}_{t_i}, \tilde{T}_{t-1}) \), is used in Eq. (10). The maximisation of the conditional probability \( p(\hat{\theta}_{t_j} | \hat{\psi}_{t_i}, \tilde{T}_{t-1}) \) now changes to that of \( p(\hat{\theta}_{t_j} | \tilde{T}_{t-1}) \) (or \( p(\hat{\psi}_{t_i} | \hat{\theta}_{t_j}, \tilde{T}_{t-1}) \)).

\[
p(\hat{\theta}_{t_j} | \hat{\psi}_{t_i}, \tilde{T}_{t-1}) = \frac{p(\hat{\theta}_{t_j} | \tilde{T}_{t-1}) p(\hat{\psi}_{t_i} | \tilde{T}_{t-1})}{\sum_i p(\hat{\theta}_{t_i} | \tilde{T}_{t-1}) p(\hat{\psi}_{t_i} | \tilde{T}_{t-1})}.
\]

Firstly, probability \( p(\hat{\theta}_{t_j} | \tilde{T}_{t-1}) \) is described using

\[
p(\hat{\theta}_{t_j} | \tilde{T}_{t-1}) = \frac{1}{N} \sum_i p(\hat{\theta}_{t_i} | \tilde{T}_{t-1}),
\]

where \( N \) is the number of the involved image points. It indicates that the posterior probability \( p(\hat{\theta}_{t_j} | \tilde{T}_{t-1}) \) can be determined by the mean of the individual joint densities \( p(\hat{\theta}_{t_i} | \hat{\psi}_{t_i}, \tilde{T}_{t-1}) \) over the selected image points.

Secondly, the densities \( p(\hat{\psi}_{t_i} | \hat{\theta}_{t_j}, \tilde{T}_{t-1}) \) are assumed to be a member of the exponential family of densities [22]. In case of the \( n \)-dimensional Gaussian \( g \), the total probability is the mixture of multinomial probabilities:

\[
p(\hat{\psi}_{t_i} | \hat{\theta}_{t_j}, \tilde{T}_{t-1}) = \sum_m g_m(\hat{\theta}_{t_j}, \tilde{T}_{t-1}) \sum_k g_k|m(\hat{\theta}_{t_j}, \tilde{T}_{t-1}) p_m(k | \hat{\psi}_{t_i}),
\]

where \( \tilde{T}_{t-1} \) is the target. Using the Markovian properties, it is expected to maximise the conditional log-likelihood as follows:

\[
Q(\tilde{T}_t | \tilde{T}_{t-1}) = \frac{1}{2\sigma^2} \sum_i \sum_j p(\hat{\theta}_{t_j} | \tilde{T}_{t-1}) (\hat{\psi}_{t_i} - \mu)^T (\hat{\psi}_{t_i} - \mu),\]

where \( g_m \) and \( g_k|m \) are two probabilities, depending on the inputs/parameters. \( \mu \) is the mean of \( \hat{\psi}_{t_i} \). Therefore, Eq. (10) can be re-written as follows:

\[
Q(\tilde{T}_t | \tilde{T}_{t-1}) = \frac{1}{2\sigma^2} \sum_i \sum_j p(\hat{\theta}_{t_j} | \tilde{T}_{t-1}) (\hat{\psi}_{t_i} - \mu)^T (\hat{\psi}_{t_i} - \mu).
\]
where $\xi$ is the total summation of other terms after using logarithms.

(2) M-step: In this step, Eq. (12) is iteratively deployed till its maximum value has been found. Alternatively, an optimal solution $T_\ast = \mathcal{M}(T_{t-1})$ is pursued so that $Q(T_\ast | T_{t-1}) \geq Q(T_{t-1} | T_{t-1})$. To obtain an optimal solution to the above equation, one can differentiate $Q(T_\ast | T_{t-1})$ in terms of the shape distance $D_{SC}$, and then set this to be zero. Therefore, an optimal shape transformation can be obtained by solving the following numerical equations:

$$\frac{\partial Q(T_\ast | T_{t-1})}{\partial D_{SC}} = 0. \tag{13}$$

The following approximation is valid:

$$p(\theta_t \mid \psi_t, \hat{T}_{t-1}) \approx K \exp \left(-\frac{(\theta_t - \bar{\psi}_t)^T \Sigma^{-1} (\theta_t - \bar{\psi}_t)}{2}\right), \tag{14}$$

where $K = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{-\frac{1}{2}}}$, and $d$ is the number of the used image points, and $\psi_t$ and $\Sigma$ are the mean and covariance of component $\theta_t$ within a small neighborhood (e.g. $5 \times 5$). These parameters allow the point matching to be executed in a finer level. According to the definition of the shape distance/contexts, such a relationship is valid: $(\bar{\theta}_t - \psi_t) \rightarrow D_{SC}$. Hence, the left hand side of Eq. (13) can be decomposed as follows:

$$\frac{\partial Q(T_\ast | T_{t-1})}{\partial D_{SC}} \approx D_{SC} \frac{\beta \Sigma^{-1} K}{2\sigma^2} \sum_i \sum_j \left[p(\theta_t \mid \psi_t, \hat{T}_{t-1})
(\xi + \beta (\bar{\psi}_t - \mu)^T (\bar{\psi}_t - \mu)) \exp \left(-\frac{1}{2}(\bar{\theta}_t - \bar{\psi}_t)^T \Sigma^{-1} (\bar{\theta}_t - \bar{\psi}_t)\right)\right]. \tag{15}$$

Eq. (15) indicates that the shape transformation fully depends on the shape distance $D_{SC}$ and the weighted summation (within the bracket). This iterative search will stop if one of the two components experiences minor changes. This also allows the similarity between the two shapes to be quantified.

C. Algorithmic flowchart

The proposed algorithm works in this way: motion trajectories are formed by calculating the Euclidean distance between the positions of the wrist (or elbow) and shoulder joints over time. The angles (flexions) between (1) the upper and lower arms and (2) the upper arm and the human trunk are also used. Then, a number of segments of a pre-defined length, extracted from the motion trajectories, is converted from a two-dimension matrix (amplitudes vs. time) to BMP images. To form continuous curves from these measurements, a spline interpolation technique can be applied.

Afterwards, the shape distance between each segment of a motion trajectory and the reference segment (a segment of “push” patterns) is calculated. To maximise the similarity measurement between the two shapes, an EM algorithm is applied to find the best registration. See Algorithm 1 for a summary of the proposed algorithm (note that this does not include a classification step). Fig. 5 illustrates the results of combining shape context and the EM algorithm. It demonstrates that the correspondences missed in Fig. 4 have been found in Fig. 5. Table IV also shows significant difference in $I_f$ values using the proposed scheme, compared to those of Table III.

Algorithm 1 Maximising the similarity of two shapes using shape context and EM algorithm.

1. Initialise the system parameters (see Section III-B).
2. For iterations $i = 1:m$ repeat
   1. Randomly extract image points from the shapes.
   2. Calculate the transformation between the shapes via Eq. (3).
   3. Compute shape distance between the two shapes using Eq. (8).
   4. Compute the conditional log-likelihood by Eq. (12).
   5. Check if Eq. (15) leads to minor changes.
   6. If a pre-defined threshold is satisfied, stop the iteration.
3. Keep the corresponding image points which have the least shape distance.
4. Conduct shape matching once again using the inliers of the image correspondences.

V. EXPERIMENTAL WORK

In this section, the proposed algorithm will be compared against several state of the art techniques for motion trajectory classification. Data acquisition is started in the first instance, followed by feature extraction. The classification procedure and a full evaluation will be introduced subsequently. Right arms have been commonly used to manipulate objects. Without loss of generality, the classification of motion trajectories of the right arms is mainly discussed herein.
A segment of 2.5 seconds is used in all the experiments reported in this paper. Such a segment may contain motion noise or instable trembles, and hence a zero-crossing method [28] is used to adaptively separate the meaningful waveforms from the backgrounds.

In this paper, the impact of segment-length variations on the classification performance is investigated. Particularly, the cases with the segment lengths of 1, 2.5 and 3 seconds are examined. To cope with the “baseline drift” problem, where the human arms cannot return to the starting positions, one can use a wavelet-based approach [21] to remove the drifts by zeroing the scaling coefficients of the discrete wavelet transform.

3) Feature extraction: Following the recommendations made in [27], the used datasets are categorised into time and frequency domains, respectively. The features in the time domain include the mean absolute value (MAV), root mean square (RMS), waveform length (WL), and slope sign changes (SSC). In the frequency domain, autoregressive coefficients of order 2 and 6 (AR2 and AR6) are used. The features of the signals are extracted from the segments with the length of 2.5 seconds and then concatenated together before they go to the classifier. The other features consist of shape distances (SD) using the classical shape contexts (SC) and the proposed scheme in this paper. Note that the amplitudes of different waveforms must have the same unit before they are picturised or stored in two dimensions.

In total, there are eight motion patterns to be classified in this section, i.e. “clapping”, “hand shaking”, “pulling”, “punching”, “pushing”, “reaching”, “slapping” and “waving”. For simplicity, these patterns are denoted as classes 1, 2, 3, 4, 5, 6, 7 and 8 and hereafter. The entire data sampling takes one hour. After a motion exercise of five minutes, there is ten minutes rest before the next practice. This interval is designed for maintaining the concentration and attention of the subjects. As an example, Fig. 8 shows the histograms of the AR6 features extracted from the eight motion patterns (the features have been averaged over the entire session).

B. Evaluation

In the evaluation, the classification results of the eight motion patterns are investigated, using the extracted features (MAV, RMS, WL, SSC, AR2, AR6, SC and proposed) with...
a classical k-means clustering method. Furthermore, the proposed algorithm is compared against the state of the art techniques, i.e., DTW [25], Levenshtein Distance on Trajectories (LDT) [13] and affine registration (AR) [14]. The first experiment is conducted comparing the performance of single feature based k-means clustering over those of different measurements (i.e., positions of the wrist and elbow joints and flexion angles). Secondly, it will be investigated how the combination of different measurements affects the classification performance. The third experiment is about the fusion of two or more different features and its application in the classification. Fourthly, the consequence of applying motion trajectories of various segment lengths to clustering approaches is examined. Afterwards, the classification performance of the proposed system is studied, when the upper limbs move at different speeds. Finally, efficiency of these algorithms is studied.

Table V: Using the extracted WL features, “clapping” and “waving” patterns obtain the highest classification rates due to their strong similarity to “pulling” in the feature space. In the meantime, “hand shaking”, “pulling”, “punching”, “pushing” and “reaching” patterns have the classification rates of approximately 70% (some variations are observed as well).

Table VI: “Slapping” and “waving” patterns only obtain the accuracy of 50% or so. This is due to the fact that their slope sign changes are less distinctive than “clapping”, “pushing” and “reaching” respectively, and the classifier cannot correctly differentiate these changes from other motion patterns. “Hand shaking” and “reaching” patterns have been successfully recognised. This success attributes to the regular SSC measurements.

Table VII: “Pushing” and “reaching” patterns have relatively lower classification rates than the remainder. This weakness is due to the fact that these two patterns share similar features with “clapping”, “reaching” and “slapping”, resulting in certain ambiguity in the classification stage.

1) Single feature based: The classification results of wrist movements using individual features MAV, RMS, WL, SSC, AR2, AR6 and SC are tabulated in Tables V-XII. In general, single feature based methods only favor some of the motion patterns in terms of classification accuracy. More details are followed.

Fig. 8. Histograms of extracted AR6 features from the eight motion patterns.

Fig. 9. Bar chart histograms of a set of WL features extracted from “clapping” and “waving” patterns, respectively.
Table VIII: “Reaching”, “slapping” and “waving” patterns cannot be effectively distinguished while the others have been successfully separated. It is because these three motion trajectories have similar features to those of “waving”, “punching” and “reaching”, respectively.

Tables IX and X show some interesting results. While these two tables demonstrate the individual roles of AR2 and AR6 features, AR2+AR6 features can help improve the classification rates if properly combined. For example, the values shown in the last three columns of these two figures show that one feature causes low classification rates, whilst the other feature augments the classification performance.

### TABLE IX
Confusion matrix of motion pattern classification using the extracted AR2 features.

<table>
<thead>
<tr>
<th>Percentages</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>66</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>100</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>84</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>52</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>74</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>78</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>81</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>47</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

### TABLE X
Confusion matrix of motion pattern classification using the extracted AR6 features.

<table>
<thead>
<tr>
<th>Percentages</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>84</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>84</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>73</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>78</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>74</td>
</tr>
</tbody>
</table>

Tables XI and XII show that, in general, the classical SC and the proposed features lead to better classification outcomes than the other features introduced above. Additionally, the classification rate of the proposed approach is 3% higher than that of the classical SC in average. To find out whether or not other measurements help obtain better classification performance than the case of only using the wrist trajectories, it is recommended to consider the motion trajectories of the elbow joint and flexion angles. Instead of showing individual results, the statistical performance of each measurement against individual features is summarised in Table XIII.

In Table XIII, DTW refers to the well established algorithm “dynamic time warping”. LDT is used to calculate the Levenshtein distance between two 3-D motion trajectories (i.e. wrist/elbow positions measured over time), and AR can be used to calculate the difference between two 2-D trajectories after the shape registration is accomplished. Fig. 10 shows the extracted DTW features, it is witnessed that the unnormalised distance of two different “clapping” segments (distance is between 10 and 20) is significantly less than that of “clapping” vs. “waving” patterns. In this case, “clapping” can be separated from “waving”.

Table XIII indicates that the proposed scheme leads to the best classification accuracy. To further compare the performance of the classical SC and the proposed approaches, the computed classification rates are studied using the Student’s t-test. Using a 0.95 confidence, we calculate the p-values for three pairs of classification rates, i.e. wrist(SC)/wrist(proposed), elbow(SC)/elbow(proposed) and flexion(SC)/flexion(proposed). Three p-values (3.92 × 10⁻⁴; 4.24 × 10⁻¹¹; 1.60 × 10⁻¹⁰) are available, which suggest that the proposed algorithm is statistically better than the classical SC approach. Fig. 11 shows that the performance of the proposed algorithm attributes to less mean shape distance (approx. 7%) than the classical one. It also shows that the proposed scheme leads to more correspondences than the classical SC.

### TABLE XI
Confusion matrix of motion pattern classification using the extracted SC features.

<table>
<thead>
<tr>
<th>Percentages</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>89</td>
<td>0</td>
<td>5</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>92</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>7</td>
<td>93</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>86</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>96</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>98</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>91</td>
<td>0</td>
</tr>
</tbody>
</table>

### TABLE XII
Confusion matrix of motion pattern classification using the proposed approach.

<table>
<thead>
<tr>
<th>Percentages</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>89</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>2</td>
<td>94</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>89</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>98</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>0</td>
<td>97</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>
TABLE XII
STATISTICS OF CLASSIFICATION PERFORMANCE USING VARIOUS MEASUREMENTS AND FEATURES: (MEAN±STANDARD DEVIATION)\%.

<table>
<thead>
<tr>
<th>Types</th>
<th>WL</th>
<th>SSC</th>
<th>RMS</th>
<th>MAV</th>
<th>AR2</th>
<th>AR6</th>
<th>SC</th>
<th>Proposed</th>
<th>DTW</th>
<th>LDT</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrist</td>
<td>80.2±13.1</td>
<td>75.9±19.7</td>
<td>79.8±27.5</td>
<td>80.4±23.6</td>
<td>79.3±15.7</td>
<td>86.9±11.7</td>
<td>91.5±4.4</td>
<td>94.6±4.1</td>
<td>88.4±7.1</td>
<td>91.2±8.7</td>
<td>89.0±9.8</td>
</tr>
<tr>
<td>Elbow</td>
<td>74.8±15.6</td>
<td>81.2±11.1</td>
<td>84.8±15.8</td>
<td>73.4±16.9</td>
<td>84.3±10.2</td>
<td>82.3±16.7</td>
<td>87.2±8.3</td>
<td>92.7±5.4</td>
<td>86.1±9.1</td>
<td>89.6±9.4</td>
<td>86.3±9.8</td>
</tr>
<tr>
<td>Flexion</td>
<td>81.8±15.8</td>
<td>85.2±8.8</td>
<td>72.1±17.5</td>
<td>79.2±11.1</td>
<td>82.7±15.9</td>
<td>81.0±13.6</td>
<td>90.3±6.4</td>
<td>93.7±3.5</td>
<td>84.0±4.4</td>
<td>–</td>
<td>90.5±6.9</td>
</tr>
</tbody>
</table>

C. Combined measurement based

For a better demonstration purpose, the case, where two out of the available three measurements are combined, is shown here. As an example, the combination of wrist and flexion measurements is used with individually extracted features for clustering. Table XIV illustrates the classification results (diagonal elements of the confusion matrix), taking into account the combination of wrist/flexion trajectories. Comparing the results shown in Tables XIV and XIII, one can observe that the combinatorial measurements leads to approx. 0-10% increments in the classification rates. This observation holds for three different combinations of the wrist/elbow/flexion measurements.

D. Combined feature based

In this sub-section, tests are performed using different combinations of features for the motion classification purpose. First of all, let us take a look at the combination of two features, e.g. RMS and MAV. Secondly, it is worthy to investigate how the combination of three features (e.g. WL, SSC and SC) affects the classification performance. The experimental details are followed.

Firstly, features RMS and MAV are extracted from one of the three measurements (wrist, elbow and flexion) for classification. Table XV denotes the classification performance against the three measurements. One observes that the combination of the two features actually helps enhance the classification rates of the three motion patterns, compared to the results shown in Table XIII. Secondly, one can concatenate three features WL, SSC and the proposed. To choose appropriate features for the classification purpose, the “FCBF” algorithm reported in [46] is applied for feature selection, where features are sorted using a relevance score. Table XVI shows significantly augmented classification rates. This is due to the fact that FCBF is capable of identifying the discriminative features in the feature pool.

TABLE XIV
CLASSIFICATION RESULTS OF COMBINING THE WRIST/FLEXION TRAJECTORIES WITH INDIVIDUAL FEATURES (MEAN WITH STANDARD DEVIATION).

<table>
<thead>
<tr>
<th>Features</th>
<th>Classification rates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WL</td>
<td>83.5±12.4</td>
</tr>
<tr>
<td>SSC</td>
<td>86.1±15.3</td>
</tr>
<tr>
<td>RMS</td>
<td>84.0±22.8</td>
</tr>
<tr>
<td>MAV</td>
<td>81.9±16.7</td>
</tr>
<tr>
<td>AR2</td>
<td>86.8±9.7</td>
</tr>
<tr>
<td>AR6</td>
<td>86.0±15.1</td>
</tr>
<tr>
<td>SC</td>
<td>92.6±5.5</td>
</tr>
<tr>
<td>Proposed</td>
<td>94.8±6.0</td>
</tr>
</tbody>
</table>

TABLE XV
CLASSIFICATION RESULTS OF COMBINING FEATURES RMS AND MAV FOR THE WRIST, ELBOW AND FLEXION MEASUREMENTS (MEAN WITH STANDARD DEVIATION).

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Classification rates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrist</td>
<td>84.6±9.7</td>
</tr>
<tr>
<td>Elbow</td>
<td>88.9±12.5</td>
</tr>
<tr>
<td>Flexion</td>
<td>86.1±13.9</td>
</tr>
</tbody>
</table>

TABLE XVI

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Classification rates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrist</td>
<td>91.6±7.5</td>
</tr>
<tr>
<td>Elbow</td>
<td>93.4±8.1</td>
</tr>
<tr>
<td>Flexion</td>
<td>90.1±6.8</td>
</tr>
</tbody>
</table>

E. Changes of segment length

This is an important aspect as changed segment lengths may cause degraded classification performance. As an example, our experimental results show how the varied segment lengths affect the final classification results. Table XVII denotes the outcomes of motion pattern classification only using the proposed algorithm with segment lengths of 1, 2.5 and 3 seconds individually. Note that this example utilises the wrist, elbow and flexion measurements independently. No significant changes in the classification results have been observed to the varied segment lengths. Similar results have also been found using other features.

F. Changes of motion speeds

Speed variations of human upper limbs are inevitable in daily life. These variations may also affect the accuracy of a classifier. To find out what will happen in different motion speeds, arm motion is performed at four different speeds (approximately): 0.1, 0.15, 0.2, and 0.25 ms\(^{-1}\). As an example, the evaluation of the proposed algorithm is here introduced, according to the experimental set-up that has been used to generate Table XIII. The percentage difference between the newly obtained classification rates and those reported in the corresponding columns of Table XIII is denoted.

TABLE XVII
CLASSIFICATION RESULTS OF COMBINING FEATURES WL, SSC AND THE PROPOSED FOR THE WRIST, ELBOW AND FLEXION MEASUREMENTS USING THE SEGMENT LENGTHS OF 1, 2.5 AND 3 SECONDS (UNIT: %).

<table>
<thead>
<tr>
<th>Measurements</th>
<th>1s</th>
<th>2.5s</th>
<th>3s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrist</td>
<td>91.5</td>
<td>95.3</td>
<td>93.6</td>
</tr>
<tr>
<td>Elbow</td>
<td>94.2</td>
<td>92.9</td>
<td>89.7</td>
</tr>
<tr>
<td>Flexion</td>
<td>90.5</td>
<td>93.5</td>
<td>95.8</td>
</tr>
</tbody>
</table>
in Fig. 12. Even though motion speeds significantly change, the percentage difference appears to be minor (<5%).

G. Computational costs

Efficiency is a critical aspect in real time processing. In this work, the computational requirements of different algorithms in the feature extraction step is investigated. The algorithms that involve in this evaluation consist of the classical SC, WL, DTW and the proposed scheme in this paper. The computational costs are calculated and then averaged over 20 iterations, which are performed in the Matlab environment (The MathWorks, Natick, USA), running on a PC with Intel Xeon Quad-Core CPU 2.4GHz and 3.0GB RAM. Table XVIII shows that the WL based algorithm is the fastest, whilst the proposed approach uses the majority of the computational efforts to search for the best registration.

VI. CONCLUSIONS AND FUTURE WORK

A new shape feature extraction scheme for motion pattern recognition has been presented in this paper. An expectation-maximisation algorithm has been used with the classical shape contexts scheme to pursue maximum similarity between two shapes. The proposed strategy has been compared to the state of the art techniques, and the experimental results favour the proposed approach in the classification of eight motion patterns of the upper limbs.

Future work will be directed towards investigating the combination of multiple features in the classification of motion trajectories. The current study has revealed that the combination of different features can result in better classification results. It will be very interesting to find out in the future study which physical feature dominates the classification in the combinatorial form, and how the best feature for reliable classification can be used consistently. On the other hand, reducing computational complexity and system requirements of the current design is also one of our objectives.

REFERENCES


Huiyu Zhou received his BEng degree in radio technology from Huazhong University of Science and Technology, China, in 1990. He was awarded an MSc degree in biomedical engineering from the University of Dundee, UK, and an PhD degree in computer vision from the Heriot-Watt University, Edinburgh, United kingdom. Currently, he holds a permanent post in ECTE, Queen’s University Belfast, United Kingdom. His research interests include computer vision, human motion analysis, intelligent systems and human computer interface.

Huishong Hu (M’94-SM’01) is a Professor in the School of Computer Science and Electronic Engineering, University of Essex, United Kingdom, where he is leading the Human Centred Robotics Group. He is the author or coauthor of more than 300 papers published in various journals, books, and conferences, and the recipient of several best paper awards. His current research interests include autonomous mobile robots, human-robot interaction, evolutionary robotics, multi-robot collaboration, embedded systems, pervasive computing, sensor integration, intelligent control, and networked robotics. He is the Editor-in-Chief of International Journal of Automation and Computing, and an Executive Editor of International Journal of Mechatronics and Automation.

Honghui Liu (M’02-SM’06) received his Ph.D degree in intelligent robotics from Kings college, University of London, UK, in 2003. He is Professor in Intelligent Systems at the University of Portsmouth. He previously held research appointments at the University of London and University of Aberdeen, and project leader appointments in large-scale industrial control and system integration industry. He is interested in approximate computation, pattern recognition, intelligent video analytics, cognitive robotics and their practical applications with an emphasis on robot-robot and robot-human interaction.

Jinsheng Tang (M’00-SM’03) received the Ph.D. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 1998. He joined ATR Media Integration and Communication Research Laboratories, Japan, as an Invited Researcher in 1998. In 2000, he joined Harvard Medical School as a Postdoctoral Researcher. In 2001, he joined the University of Virginia, and was there for about three years. He was a Visiting Fellow at the National Cancer Institute, National Institutes of Health (Bethesda, MD), and as a Senior Engineer in Intel, Hudson, MA. From 2006 to 2010, he was an Assistant Professor at Alcorn State University, Alcorn State, MS. He is now an Associate Professor of Michigan Technological University, United States. Dr. Tang published more than 80 Journal and conference papers on image processing and medical imaging. He is a senior member of IEEE and a member of the Technical Committee on Information Assurance and Intelligent Multimedia-Mobile Communications, IEEE SMC society.