ISSN 1751-9659
Received on 17th April 2014
Revised on 27th January 2015
Accepted on 21st May 2015
doi: 10.1049/iet-ipr.2014.0622
www.ietdl.org

Topomorphological approach to automatic posture recognition in ballet dance

Sriparna Saha ✉, Amit Konar

Artificial Intelligence Laboratory, Electronics and Tele-Communication Engineering Department, Jadavpur University, Kolkata, India
✉ *E-mail: sahasriparna@gmail.com*

Abstract: The proposed system aims at automatic identification of an unknown dance posture referring to the 20 primitive postures of ballet, simultaneously measuring the proximity of an unknown dance posture to a known primitive. A simple and novel six stage algorithm achieves the desired objective. Skin colour segmentation is performed on the dance postures, the output of which is dilated and is processed to generate skeletons of the original postures. The stick figure diagrams laden with minor irregularities are transubstantiated to generate their affirming minimised skeletons. Each of the 20 postures based on their corresponding Euler number are categorised into five groups. Simultaneously the line integral plots of the dance primitives are determined by performing Radon transform on the minimised skeletons. The line integral plots of the fundamental postures along with their Euler number populate the initial database. The group of an unknown posture is determined based on its Euler number, while successively the unknown posture's line integral plot is compared with the line integral plots of the postures belonging to that group. An empirically determined threshold finally decides on the correctness of the performed posture. While recognising unknown postures, the proposed system registers an overall accuracy of 91.35%.

1 Introduction

Ballet is an artistic dance form, performed to music using unequivocal and highly formalised set steps and kinesics. The proposed body of work deals with automatic posture recognition of the various dance sequences of ballet. Recognition, modelling and analysis of the various postures comprising the dance form help an enthusiast to learn the art by intelligibly interacting with the proposed system. The motivation behind this novel work is to propose a system for automatic learning of ballet dance in a simple and flexible manner.

Substantial work has been done in the field of posture recognition; however, very few of them deal with posture recognition in specific dance procedures. The work discussed in [1] deals with the sequencing of dance postures using stick figure diagrams. However, the dependence of the algorithm on motion sensor devices escalates the cost of the interactive setup. In addition to that, the algorithm deals exclusively with the leg movements of the performer and in the process neglects the specifications of the other body parts. Without considering the entire posture of a performer, capturing the essence of a dance form as elaborate as ballet is not possible.

Guo and Qian [2] deals with three-dimensional (3D) dance gesture recognition. While comparing the algorithm proposed in this paper to the one presented by Guo and Qian [2], we find that the premise for comparison is quite different. Guo and Qian have used 3D images and have used a train and test relevance vector machine. The algorithm proposed in this paper deal with 2D posture recognition of ballet dance postures and uses a simple camera instead of the dual camera setup that has been used in [2] for orthogonal projections. The algorithm proposed by Guo and Qian registers a classification accuracy of 81.9%, while the algorithm proposed in this paper registers a classification accuracy of 91.35%. The proposed algorithm successfully tackles the anomalies created because of different body shapes of ballet dancers that Guo and Qian fail to address.

The authors in [3] proposed a dance posture recognition technique with the help of which issues related to the texture of the dress, distance of the performer with respect to the camera and body structure of the dancer is effectively addressed. With the help of the proposed algorithm, the dancer is effectively extracted from the background in which (s)he is performing. Apart from the above-mentioned procedures dealing exclusively with dance postures, there exists a comprehensive body of work dealing entirely with human motion analysis and posture recognition.

Analysis of human gesture from different perspectives is an important aspect. These can be done using image sequence analysis [4] also by representing the human body using stick figure diagrams [5]. In [6], 2D posture recognition is done using silhouette extraction. Here human body is segmented from a complex background using a statistical metod, and matching of postures is performed using genetic algorithm. For the purpose of designing an alarm system for elderly healthcare in [7], the authors have used support vector machine classification after segmenting the object from the background using background difference methods and extracting features in terms of shape and area. In [8], activity monitoring of elders is proposed with the help of silhouette extraction. In both [7, 8], the work involves sequences of images, that is, postures are recognised from video.

Two similar approaches have been proposed in [9, 10]. Here to approximate the curved lines of the skeletons by straight lines using chain codes. Drawback of both the two papers lies in restricting the RGB image of the dancer needs to be centralised. This factor unfortunately is of important concern in learning ballet dance. In this proposed work, we modify the line integral plots by rejecting leading and trailing zeroes and thus the problem of centralising does not appear. Also the accuracy of the proposed work is better than those of [9, 10]. This proposed work deals with static posture recognition and thus cannot be compared with dynamic gesture recognition techniques like [11, 12].

In this work, we are dealing with 20 basic postures of ballet dance. RGB colour images pertaining to the primitive postures of ballet constitute the initial database. Unclassified ballet postures act as inputs to the algorithm. Skin colour segmentation is performed on the input images, to get rid of the unnecessary background information. The extracted postures are modelled using skeletonisation. Information derived from the stick figure diagrams include the Euler number of the skeleton and the line integral plot using Radon transform. Finally, a sum of correlation coefficients measure is employed to classify the unknown dance posture. The information extracted is processed according to the proposed technique, which not only detects a posture but additionally determines its correctness. The preferred system makes learning ballet an interactive exercise whereby the learner interacts with the computer in an efficient as well as economical manner.

2 Methodology

This section explains the pre-processing steps and the principles used for posture recognition in great detail. The flowchart corresponding to the proposed algorithm is given in Fig. 1.

2.1 Morphological processing

Colour images of the ballet postures in RGB format act as inputs to the proposed algorithm. A particular dress pattern is recommended while performing ballet and owing to its specifications, a major portion of the dancer's body remains exposed. Hence, skin colour segmentation is used to extract the performer from the background. The skin and non-skin pixels are distinguished from each other using unimodal Gaussian model [13]. Let $\boldsymbol{m}$ is the mean vector and $\boldsymbol{C}$ is the co-variance matrix of skin colour. Again, x is a skin pixel if the following equation holds

$$(x - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{C}^{-1} (x - \boldsymbol{m}) - \left(x - \boldsymbol{m}_n\right)^{\mathrm{T}} \boldsymbol{C}_n^{-1} (x - \boldsymbol{m}_n) \le \tau \qquad (1)$$

where m_n and C_n are the mean and co-variance of skin and non-skin colours, respectively. For our purpose, we have taken threshold value (τ) as 0.01 [14].

On performing skin colour segmentation, certain sections of the original RGB image are wrongly classified as skin segments. In order to get rid of this problem, connected components are considered in the binary image and components having less than 100 pixels (each connected to its eight neighbours) are automatically removed from it.

The skin colour segmented images are dilated with the help of two line elements, each of length 3 units and corresponding to angles 0° and 90°, respectively. The dilation procedure corrects the minor irregularities present in the segmented image to a considerable extent.

In the next step, the dilated images are skeletonised giving rise to stick figure representations of the concerned dance postures. The skeletons so formed contain certain irregular constructs, which add no valuable information to the image. As a result of which shorter length lines are eliminated from the medial axis images with the help of a morphological spur operation. The minimised skeletons finally contain the significant lines that maximally determine a dance posture. Fig. 2 contains a step-by-step depiction of the entire process.

2.2 Euler number-based grouping of postures

A total of 20 different postures are grouped into five different categories based on the Euler number [15] of the minimised skeletons of the corresponding postures. The generated skeletons contain the variable number of open-loop lines and closed-loop holes. The Euler number (E) is determined after calculating the number of holes (H) and connected components (CCs) in a particular skeleton.

$$E = \mathrm{CC} - H \qquad (2)$$

The Euler numbers range from −1 to +3 and based on the number of CCs and holes present in the postures, they are categorised into individual groups (G). Table 1 presents the five different categories and the postures belonging to them, whereas Fig. 3 provides examples of postures belonging to each category.

2.3 Radon transform

Skeletons belonging to the same group are differentiated with the help of Radon transform [16, 17]. Each pixel constituting the minimised skeletons is projected along the x-axis and the y-axis. Fig. 4 represents the Radon transform, wherein the red colour represents projections along the x-axis (i.e. along 0° angle) and the blue colour represents projections along the y-axis (i.e. along 90° angle), respectively. Considering the image in the spatial domain (x, y) mapped along the projection domain (p, θ), the required Radon transform of a function $f(x, y)$ denoted by $R(p, \theta)$ can be

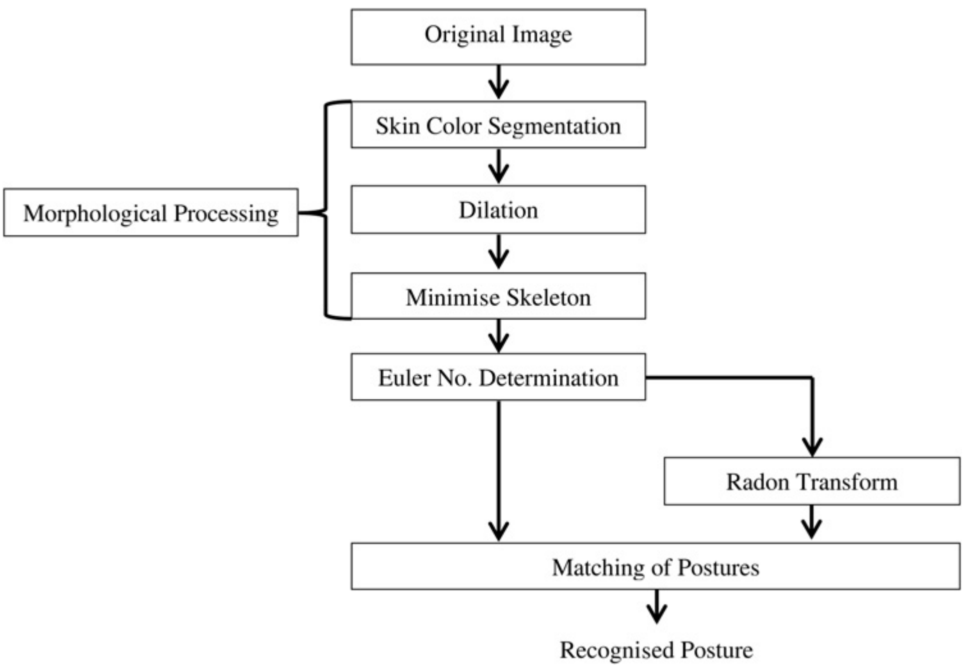


Fig. 1 *Block diagram of the proposed algorithm*

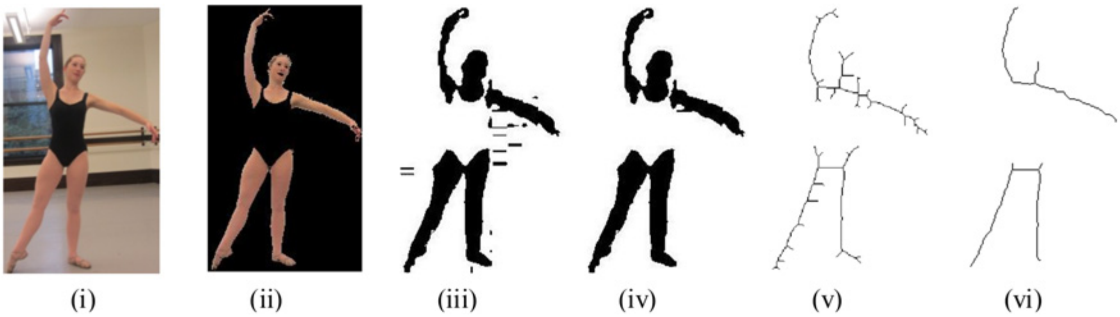


Fig. 2 *Ecarte devant posture:*

(i) Original RGB image
(ii) Skin colour segmented image
(iii) Binary representation of segmented image
(iv) Removal of unwanted pixels
(v) Morphological skeleton image
(vi) Spurring of the skeleton image to remove lines less than specified length 16 [inverted images are shown for (iii)–(vi) for better visualisation]

Table 1 Classification of fundamental postures into five groups

Euler no.	Groups (G)	Example
−1	1	arms first , posture front
0	2	arms third, arms fourth, releve
1	3	arms fifth, posture side
2	4	attitude front, croise derriere, croise devant, ecarte devant, efface devant, en face
3	5	arabesque, arms second, attitude back, attitude side

expressed as

$$
\begin{aligned}
R(p,\,\theta) &= \int_{-\infty}^{+\infty} f(x,\,y)\,\mathrm{d}q \\
&= \int_{-\infty}^{+\infty} f(p\cos\theta - q\sin\theta,\,p\sin\theta + q\cos\theta)\,\mathrm{d}q'
\end{aligned} \tag{3}
$$

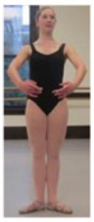
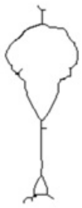
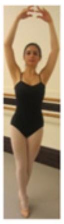
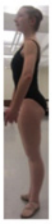
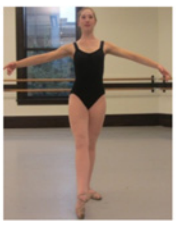
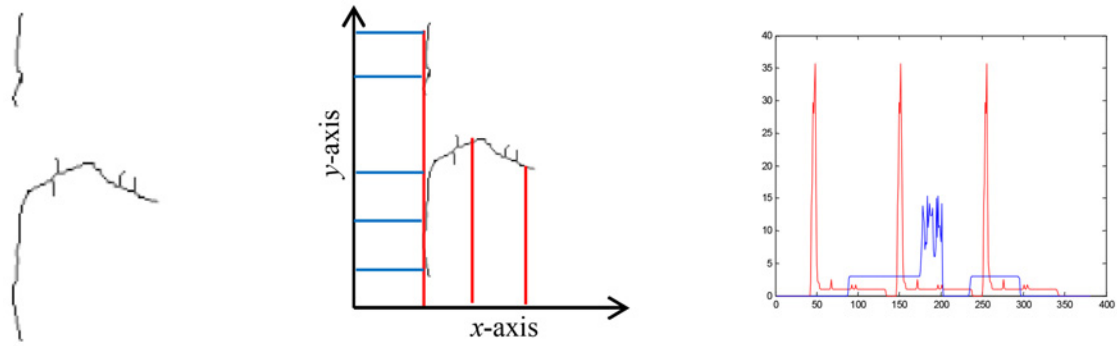
Posture Name	RGB Image	Minimised Skeleton	Feature	Group No.
Arms First			1 − 2= − 1	1
Releve			2 − 2=0	2
Posture Side			1 − 0=1	3
En Face			2 − 0=2	4

Fig. 3 *Group wise examples of relevant postures*

IET Image Process., 2015, Vol. 9, Iss. 11, pp. 1002–1011



(i) (ii) (iii)

Fig. 4 *Attitude front posture:*

(i) Minimised skeleton image
(ii) Projection along the x-axis (shown by the red line) and the y-axis (shown by the blue line)
(iii) Line integral plot

2.4 Matching of postures

The final stage of the proposed algorithm deals with recognition of an unknown ballet posture with the help of matching of their corresponding Radon transforms. An initial database is prepared by storing the radon transforms of the 20 basic dance postures of ballet.

The line integral plot of an unknown posture is determined and is matched with the line integral plots of all the postures belonging to the group in which the unknown posture belongs. To nullify the effect of height, weight and body structure of the dancer the line integral plot of the unknown and known postures are scaled accordingly. For the purpose of scaling, the leading and trailing zeroes which often form a part of the line integral plot in the two axes are neglected and the pertinent portion of the plot is considered. This approach helps us to compare non-centralised images, adding a lot of flexibility to the entire procedure. The number of points from relevant portions are either scaled up or scaled down to ensure that the number of such points is a multiple of 50. In order to scale down, the number of points that needs to be deleted is selected at equal intervals. In order to scale up, the number of points that needs to be appended is added at regular intervals. The new points added at regular intervals simply copy the existing value of its preceding point. The mathematical formulations shown below represent the scaling operation.

$$\lambda(l) = \begin{array}{l} 50-(l(\bmod 50)) \text{if} \, l(\bmod 50) > 25 \\ l(\bmod 50) \text{if} \, l(\bmod 50) \leq 25 \end{array} \tag{4}$$

where l refers to the number of points in the line integral plot except the ones corresponding to the leading and trailing zeroes. λ refers to the number of points that need to be added or subtracted from the existing line integral plot. As the function states, the number of points that needs to be added or subtracted depends on the l(mod 50) value. If the value is less than or equal to 25, then, that many points are deleted from the line integral plot. The points are chosen at regular interval and in general do not alter the plot to any considerable extent. If the value is greater than 25, 50−(l(mod 50)) points are added to the plot. These points are added at regular intervals and it replicates the value of its preceding point. After scaling, 50 points each sampled at regular intervals from both the axes are considered for both the unknown and known postures. The sampling frequency samp_rate is given as

$$\text{samp\_rate} = \frac{(l+(/-)\lambda)}{50} \tag{5}$$

Once we get 50 such points and their corresponding line integral plots with respect to both the x-axis and the y-axis we compare the unknown and known postures. Correlation coefficients [18] corresponding to both the x-axis and the y-axis are calculated for each such pair of postures. The correlation coefficient calculated along the x-axis is given as

$$Corr_x = \frac{\sum_{m_x}\left(A_{m_x}-\bar{A}\right)\left(B_{m_x}-\bar{B}\right)}{\sqrt{\left(\sum_{m_x}\left(A_{m_x}-\bar{A}\right)^2\right)\left(\sum_{m_x}\left(B_{m_x}-\bar{B}\right)^2\right)}} \tag{6}$$

While the correlation coefficient calculated along the y-axis is given as

$$Corr_y = \frac{\sum_{m_y}\left(A_{m_y}-\bar{A}\right)\left(B_{m_y}-\bar{B}\right)}{\sqrt{\left(\sum_{m_y}\left(A_{m_y}-\bar{A}\right)^2\right)\left(\sum_{m_y}\left(B_{m_y}-\bar{B}\right)^2\right)}} \tag{7}$$

Here m varies from 1 to 50 for both (6) and (7) and the correlation coefficient is calculated between the sample points present in the x-axis and the y-axis corresponding to the unknown posture and one of the known postures belonging to the same group as suggested by the Euler number of the unknown posture.

If the summation of the correlation coefficients is greater than a pre-determined threshold, the unknown posture is classified as the primitive posture having the highest sum of correlation coefficient value. The summation of the correlation coefficients is given as

$$Corr_{sum} = Corr_x + Corr_y \tag{8}$$

It must be greater than the empirically determined threshold τ to qualify as a valid identified posture. The threshold value (τ) for the proposed algorithm is 1.0 and is calculated empirically. The unknown posture is identified as the primitive posture in the same group having maximum Corr_{sum} value. Fig. 5 corresponds to the scaling and sampling procedure where the black stars in Fig. 5(iii) denoted the new points that are added to the plot. If the sum of correlation coefficients of the unknown posture and known posture pairs is less than the pre-determined threshold, the unknown posture is classified as a non-dance posture.

3 Experimental results

The performance of the proposed algorithm is measured in this section and it contains an exemplified detailed analysis of the classification procedure for both dance and non-dance postures.

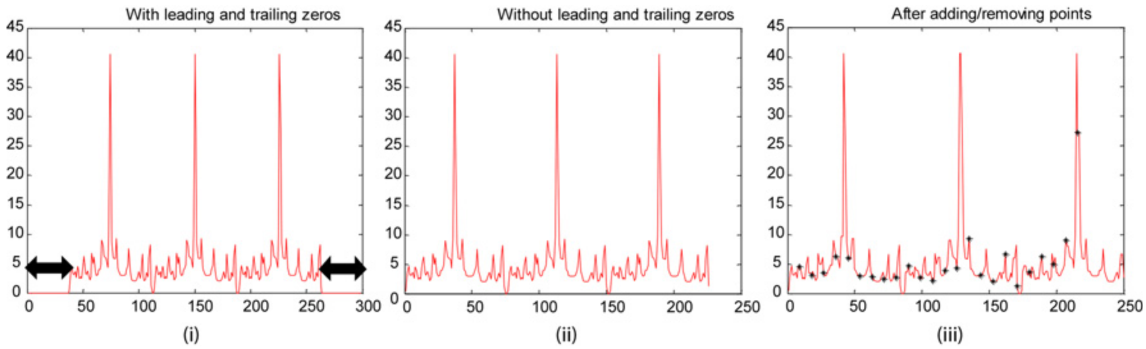


Fig. 5 *Scaling and sampling procedure*

(i) Line integral plot showing the leading and trailing zeroes
(ii) Line integral plot ignoring the leading and trailing zeroes
(iii) Scaling up, black star shows the points added

3.1 Recognition of unknown posture

The Euler number of the unknown posture is 2 and it corresponds to the fourth group as shown in Fig. 6. On determination of the Euler number, Radon transform is performed on the minimised skeleton of the unknown posture. The line integral plot of the unknown posture is compared with the line integral plot of the six postures that belong to group 4. The unknown posture is recognised as the posture in the fourth group corresponding to the maximum sum of correlation coefficients. According to the defined computations, the unknown posture is recognised as 'croise derriere' which being correct, necessarily justifies the proposed algorithm. Fig. 6 presents a step-by-step comparison of the unknown posture and posture 'croise derriere'. The initial line integral plot of the unknown posture had 539 points both along the *x*-axis and the *y*-axis. However, after ignoring the leading and trailing zeroes we arrived at 412 and 204 points along the *x*-axis and the *y*-axis, respectively. Twelve and four points are deleted at regular intervals to scale down the plot, respectively. Finally, 50 points are sampled along the *x*-axis and the *y*-axis, respectively. The same sets of plots are illustrated for posture 'croise derriere' with which the unknown posture is recognised. The known posture had 311 and 261 points along the *x*-axis and the *y*-axis, respectively. In a similar way, 50 points are sampled in the *x*-axis and the *y*-axis, respectively. Once an equal set of points and their values are determined, the sum of correlation coefficient is determined.

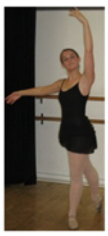
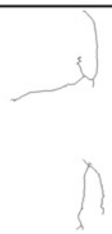
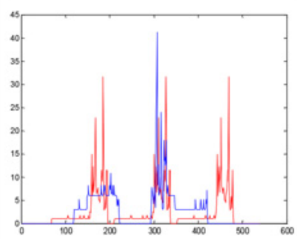
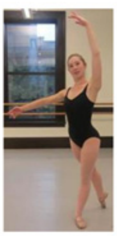
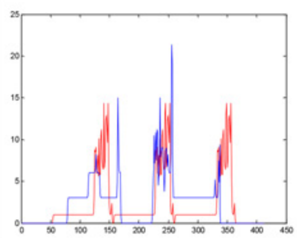
Attributes	RGB Image	Minimum Skeleton	Euler No.	Group No.	Radon Transform
Unknown Posture			2	4	
Known 'Croise Derriere'			2	4	

Fig. 6 *Comparison between unknown posture and known 'croise derriere'*

Table 2 Comparison between unknown posture and postures of group number 4

	Known postures					
	Names	$Corr_x$	$Corr_y$	$Corr_{sum}$	$>\tau$?	Index of best matched posture
unknown posture	attitude front	−0.10724	−0.4638	−0.57104	no	9
	croise derriere	0.82216	0.61589	1.43805	yes	
	croise devant	0.47912	0.53202	1.0111	yes	
	ecarte devant	−0.21399	0.022809	−0.19118	no	
	efface devant	0.36038	0.37146	0.73184	no	
	en face	−0.13536	−0.23761	−0.37297	no	

[?] in [$>\tau$?] is given to denote that whether the $Corr_{sum}$ value is greater than τ or not

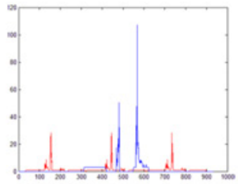
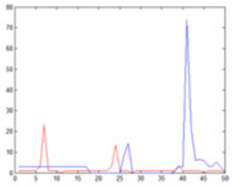
Posture Type	Unknown Posture	RGB image of unknown posture	Minimised skeleton	Line integral plot	Sampled line integral plot
					
Known Postures	Names	$Corr_x$	$Corr_y$	$Corr_{sum}$	$>\tau$?
	Arabesque	-0.0271	0.271	0.244	No
	Arms Second	-0.0122	0.259	0.247	No
	Attitude Back	-0.0542	-0.028	-0.082	No
	Attitude Side	-0.0605	0.307	0.246	No

Fig. 7 *Comparison between unknown posture and postures of group 5*

Table 2 presents the correlation coefficients corresponding to the x-axis and the y-axis of the sampled plot of the unknown posture and the known postures present in group 4. Now here $\mathrm{Corr}_{\mathrm{sum}}$ is greater than τ for more than one posture, so the posture with the highest sum of correlation coefficient value is taken as the result. The $\mathrm{Corr}_{\mathrm{sum}}$ values of unknown posture when comparing with croise derriere and croise devant is obtained as 1.43805 and 1.0111, respectively. As 1.43805 is greater than 1.0111, thus the result is posture 'croise derriere' with $\mathrm{Corr}_{\mathrm{sum}}$ as 1.43805. This value is way above the threshold and is the maximum value corresponding to group 4 for unknown posture.

An unknown posture when fails to register a sum of correlation coefficient value greater than 1.0 with any posture belonging to its group, is recognised as a non-dance posture. Fig. 7 depicts a non-dance posture using RGB image, minimised skeleton, line integral plot and also scaled and sampled plot. The sampled plot is compared with sampled plots of the postures present in group 5 (the Euler number of the unknown posture is 3). Fig. 7 also contains the sum of correlation values calculated with respect to the unknown posture and all the postures present in group 5 and none of them is above the predetermined threshold $\tau = 1.0$. Hence, the unknown posture is classified as a non-dance posture. Thus the proposed algorithm successfully tackles the problem of false acceptance.

3.2 Determining correctness of a performed posture

The same algorithm is used to determine the correctness of a particular posture. Here determination of the group with the help of Euler number is irrelevant as the posture with which the unknown posture is supposed to be matched is known a priori. The line integral plots of both the unknown posture and 'posture front' are considered. They are scaled and sampled in the same way as discussed earlier. Fig. 8 presents a step-by-step comparison of the unknown posture and posture 'Posture Front'. The correlation coefficient of the unknown posture calculated along the x-axis with posture front is 0.76553 and that calculated along

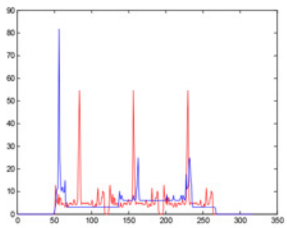
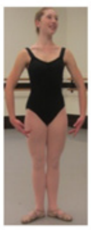
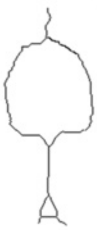
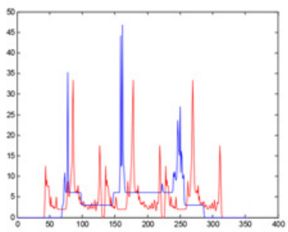
Attributes	RGB Image	Minimum Skeleton	Euler No.	Group No.	Radon Transform
Unknown Posture			-1	2	
Known 'Posture Front'			-1	2	

Fig. 8 *Comparison between unknown posture and known 'posture front'*

Frame Number	RGB Image	Recognised Posture
341		En Face
351		
362		

Fig. 9 *Results from unknown video sequence*

the y-axis is 0.26026. The summation of the correlation coefficient values is 1.0258 being greater than the threshold $\tau = 1.0$ is adjudged as a correct posture.

3.3 Determining unknown postures from video sequence

The proposed work is mainly dedicated to static 2D posture recognition from images, but can be successfully implemented in posture recognition from video sequences [19]. We have acquired a ballet video from popular website you tube (https://www.youtube.com/watch?v=b3bawTEPLtA) and break it into frames using Matlab R2013b. Now, whenever we have to recognise a posture from a specific frame, we need to capture the whole body gesture, not a part of it. The video with which we are dealing with contains several frames where the whole body of the dancer is not visible, thus we have neglected those frames. And also it is not feasible to provide results for all the 900 frames present in the video, thus we are only giving results for a few frames in Fig. 9.

Table 3 Performance evaluation of proposed algorithm

Posture name	No. of postures taken	No. of correctly identified posture	Individual recognition rate	Group-wise recognition rate
arms first	7	5	71.4286	80.1587
posture front	9	8	88.8889	
arms third	9	9	100.0000	100.0000
arms fourth	10	10	100.0000	
releve	7	7	100.0000	
arms fifth	9	8	88.8889	86.1111
posture side	6	5	83.3333	
attitude front	6	6	100.0000	89.3386
croise derriere	5	4	80.0000	
croise devant	4	4	100.0000	
ecarte derriere	4	3	75.0000	
ecarte devant	6	5	83.3333	
efface derriere	5	4	80.0000	
efface devant	7	6	85.7143	
en face	8	8	100.0000	
epaule	4	4	100.0000	
arabesque	6	6	100.0000	94.9495
arms second	9	8	88.8889	
attitude back	9	9	100.0000	
attitude side	11	10	90.9091	
non-dance posture	67	61	91.0448	91.0448
overall	208	190	91.3462	

3.4 Performance analysis

The performance metrics include precision, recall, accuracy and F1_Score. If the true positive, true negative, false positive and false negative samples are denoted by TP, TN, FP and FN, respectively.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{11}$$

$$\text{F1\_Score} = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

$$\text{Recognitionrate} = \text{Accuracy} \times 100\% \tag{13}$$

The overall performance of the proposed algorithm is presented in Table 3. The recognition rate is high when a group contains less no of primitive postures with dissimilar Radon transforms. Group 1 consists of only two postures, but the Radon transform of the

IET Image Process., 2015, Vol. 9, Iss. 11, pp. 1002–1011

Table 4 Comparison of parameters for different algorithms

Parameters	Proposed algorithm	Algorithm in [9]	Algorithm in [10]	SVM	*k*NN	LMA-NN	FCM	EDT
recall	0.9513	0.8922	0.7390	0.8488	0.8790	0.8148	0.7951	0.8382
precision	0.8859	0.7780	0.7766	0.9164	0.8660	0.7506	0.8112	0.7839
accuracy	0.9135	0.8663	0.7619	0.7903	0.8696	0.8941	0.7349	0.7239
F1_score	0.9471	0.9189	0.8789	0.8577	0.9143	0.8931	0.8900	0.8075
computation time, ms	2.6527	2.7264	2.4567	3.2951	2.9874	10.3644	3.1601	16.0215

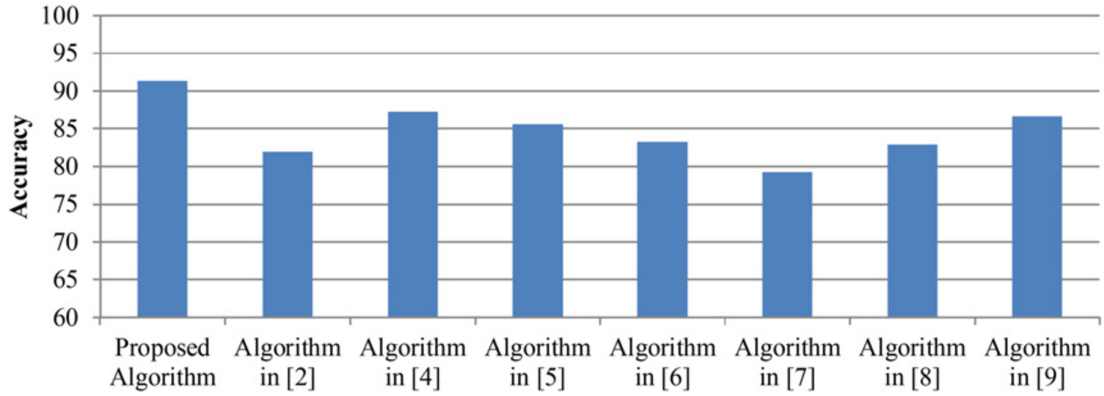


Fig. 10 *Comparison with existing posture recognition techniques*

two postures are nearly similar, so the true recognition rate falls down. However, groups 2 and 3 are composed of three and two postures, respectively, with different Radon transforms. Thus the recognition accuracy goes up for such instances. Three postures belong to group 5 with distinctly different Radon transforms, so the recognition accuracy again goes up for all such instances. The overall true recognition rate is 91.35% where 190 out of 208 unknown postures are classified correctly. The algorithm registers recognition accuracy of 80.16, 100, 86.11, 89.34, 94.95 and 91.04% for unknown postures belonging to groups 1–5 and non-dance postures, respectively. The overall confusion matrix is formed using TP = 0.9513, FN = 0.0478, FP = 0.1225 and TN = 0.8775.

Here, we have extracted mainly two types of features the Euler number and the line integral plot (after ignoring leading and trailing zeroes and applying sampling) from each image. These two features are combined together to form the feature space. This is the backbone of our proposed work. However, if only Radon transform is applied on the minimised skeleton images (i.e. if we do not group the postures based on the Euler number and without making any modifications on the line integral plots like discarding leading and trailing zeroes), then the line integral plots generated can also be treated as features. Moreover, these features can be used to examine the performance of the proposed work with a comparative framework which includes support vector machine (SVM) classifier [20], *k*-nearest neighbour (*k*NN) classification [21], Levenberg–Marquardt algorithm induced neural network (LMA-NN) [22], fuzzy *C* means clustering (FCM) [23] and ensemble decision tree (EDT) [24], algorithm in [9, 10]. The parameters of all the other competing classifiers are tuned by noting the best performances after experimental trials. SVM has been used with a radial basis function kernel whose kernel parameter has a value 1 and the classifier is tuned with a cost value of 100. The performance of *k*NN has been reported for $k = 5$ using the Euclidean distance as the similarity measure and majority voting to determine the class of the test samples. For

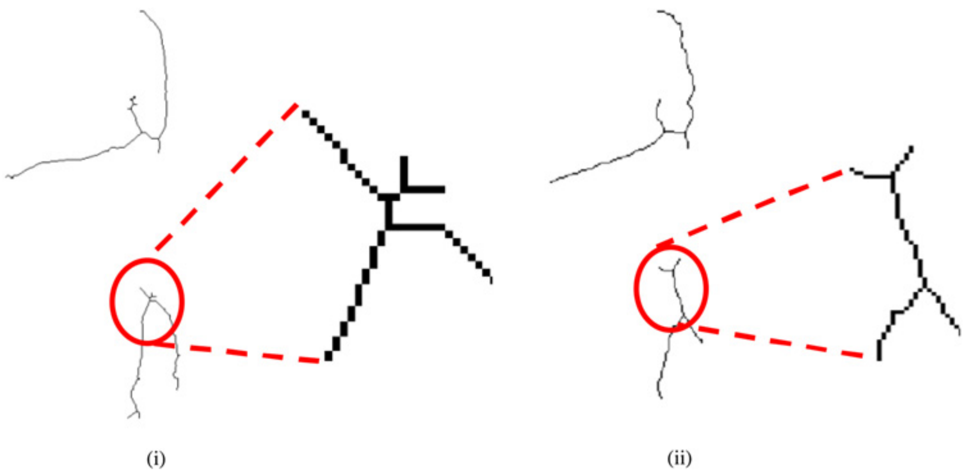


Fig. 11 *Minimised skeleton image for*

(i) Unknown
(ii) Known images from Fig. 6

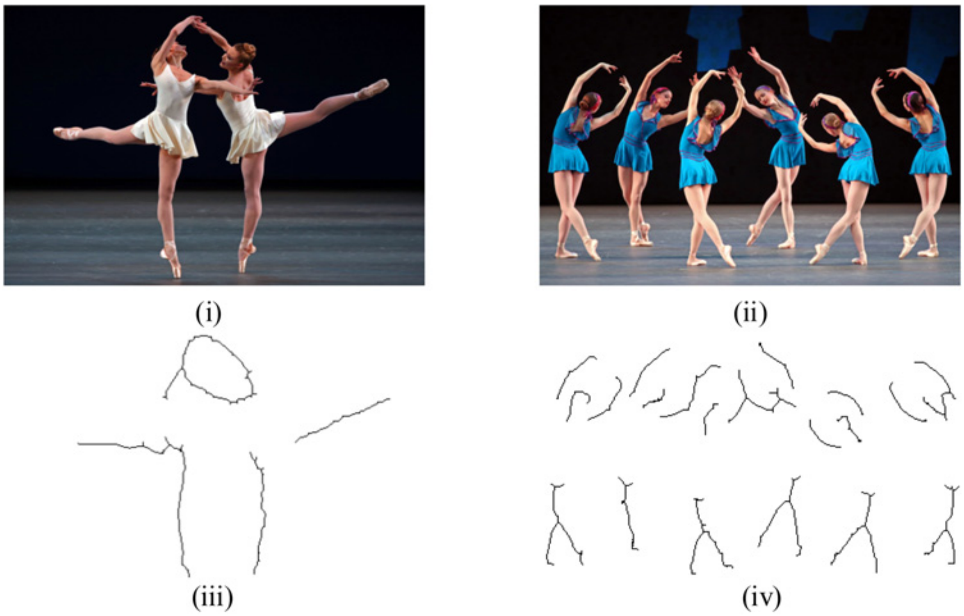


Fig. 12 *RGB images where more than one person is dancing, the main problem here is that we cannot relate the upper body of the dancer to its lower body in the skeleton images*

(i) and (ii) RGB images for more than one dancer
(iii) and (iv) Minimise skeletons for respective images

LMA-NN, the number of neurons in the intermediate layer is taken as 10, the value of the blending factor between gradient descent and quadratic learning as 0.01, the increase and decrease factors of the blending factor as 10 and 0.1, respectively, and the stopping condition is taken as the attainment of minimum error gradient value of 1×10^{-6}. EDT classifier is used based on the principle of adaptive boosting taking maximum iterations as 100. The average computation time for each image is calculated in Intel Core i3 Processor with non-optimised MatlabR2013b implementation. The values of the performance metrics of the algorithms are presented in Table 4. It is evident from Table 4 that our proposed algorithm provides best results in all parametric measures stated in (9)–(13) and also in computation time.

A comparative evaluation of the proposed procedure is difficult because of the non-availability of a standardised database. However, here we have given an assessment of the proposed work with several existing works widely varying from dance to healthcare application areas (Fig. 10).

The proposed algorithm uses the Euler number and Radon transform for feature extraction from skeleton images of dance postures. There is another way of extracting features from skeleton images using shape matching techniques. These shape detection procedures [25] are based on mainly curvature [26], similarity measures based on bottle neck distance metric [27] and turning function distance [28]. The accuracy obtained using these processes drop to 82.61, 86.35 and 79.27%, respectively. This degradation is because of the imperfections present in the skeleton images. To elaborate it, we are taking the minimised skeletons from Fig. 6. In Fig. 11, the zooming of a particular portion of the minimised skeleton images indicates the difference between the skeletons from each other when they are broken into pixel form. Thus, it is evident that Radon transform is the best choice for posture recognition in ballet dance.

The algorithm finds a concrete place for posture recognition for single dancer performing ballet. However, the proposed work cannot be implemented for more than one dancer as it is not possible to identify the dancers separately. Figs. 12(i) and (ii) explain two types of RGB images where more than one person is dancing, the main problem here is that we cannot relate the upper body of the dancer to its lower body in the skeleton images Figs. 12(iii) and (iv). If we can overcome this difficulty, then the proposed algorithm can be applied to recognise ballet dance postures where multiple persons are dancing.

3.5 McNemar's statistical test

Let f_A and f_B be two classification algorithms, both the algorithms having a common training set R. Let n_{01} be the number of examples misclassified by f_A but not by f_B, and n_{10} be the number of examples misclassified by f_B but not by f_A. Then, under the null hypothesis that both algorithms have the same error rate, the McNemar's statistic Z follows a χ^2-distribution with a degree of freedom equal to 1 [29].

$$Z = \frac{\left(|n_{01} - n_{10}| - 1\right)^2}{n_{01} + n_{10}} \tag{14}$$

Let f_A be the proposed algorithm and f_B be one of the other five algorithms. In Table 5, the null hypothesis has been rejected, if $Z > 3.84$, where 3.84 is the critical value of the χ^2-distribution for 1 degree of freedom at probability of 0.05.

Table 5 Performance analysis using McNemar's test

Competitor algorithm = f_B	Control algorithm f_A = proposed algorithm			
	n_{01}	n_{10}	Z	Comment
Algorithm in [9]	1	16	11.5294	reject
Algorithm in [10]	3	18	9.3333	reject
SVM	12	22	2.3824	accept
*k*NN	8	20	4.3214	reject
LMA-NN	6	34	18.2250	reject
FCM	4	42	29.7609	reject
EDT	3	43	33.0652	reject

4 Conclusion

Not much work has been done in the field of posture recognition of ballet. The proposed system deals with 20 fundamental dance primitives and registers an overall recognition rate of 91.35%. The procedure stated here uniquely deals with a wide range of critical dance postures. The entire endeavour proves cost effective as a single static camera produces the necessary input images for the proposed algorithm. The proposed algorithm is independent of the body type, height and weight of the ballet dancer, and hence provides even more flexibility to the learning process. The input images need not be centralised as well. The algorithm simultaneously addresses the problem of posture recognition and determination of correctness of a particular posture, thereby enhancing the effectiveness of the proposed procedure. Considering the complexity of the dance postures, an average computation time of 2.6527 s in an Intel Core i3 processor running Matlab R2013b is highly effective when compared with other standard pattern recognition algorithms.

However, certain shortcomings still do exist. The dress of the ballet dancer and the background in which (s)he is performing needs to be selected carefully, otherwise the skin colour segmentation portion of the proposed algorithm may produce sub-optimal results. The proposed algorithm is not rotation invariant; therefore the danseuse must perform in a plane which is parallel to the axis of the camera. Moreover, the performance of the algorithm drops in case of nearly identical postures, and the proposed algorithm in some cases fails to differentiate between them. These insufficiencies provide us with a lot of scope for further improvement over the proposed algorithm.

The major pre-processing for this proposed work involves skin colour segmentation. Thus, whenever we can segment out the dancer from the background using any technique (whether it is colour-based process or not), we can apply this algorithm. So to efficiently use this algorithm in other dance forms, we need to keep in mind this segmentation procedure.

In a nutshell, the system proposed for posture recognition of ballet dance may be considered as a relatively unexplored application area, and the proposed system is an attempt to address the problem with reasonable accuracy and scopes for further research.

5 References

1 LaViers, A., Chen, Y., Belta, C., Egerstedt, M.: 'Automatic sequencing of ballet poses', *IEEE Robot. Autom. Mag.*, 2011, **18**, (3), pp. 87–95

2 Guo, F., Qian, G.: 'Dance posture recognition using wide-baseline orthogonal stereo cameras'. Seventh Int. Conf. on Automatic Face and Gesture Recognition, FGR 2006, 2006, pp. 481–486

3 Silapasuphakornwong, P., Phimoltares, S., Lursinsap, C., Hansuebsai, A.: 'Posture recognition invariant to background, cloth textures, body size, and camera distance using morphological geometry'. Int. Conf. on Machine Learning and Cybernetics (ICMLC, 2010), vol. 3, 2010, pp. 1130–1135

4 Akita, K.: 'Image sequence analysis of real world human motion', *Pattern Recognit.*, 1984, **17**, (1), pp. 73–83

5 Guo, Y., Xu, G., Tsuji, S.: 'Understanding human motion patterns'. Pattern Recognition, 1994, Proc. of the 12th IAPR Int. Conf. on Vol. 2-Conf. B: Computer Vision & Image Processing, 1994, vol. 2, pp. 325–329

6 Hu, C., Yu, Q., Li, Y., Ma, S.: 'Extraction of parametric human model for posture recognition using genetic algorithm'. Proc. Fourth IEEE Int. Conf. on Automatic Face and Gesture Recognition, 2000, pp. 518–523

7 Feng, G., Lin, Q.: 'Design of Elder alarm system based on body posture reorganization'. Int. Conf. on Anti-Counterfeiting Security and Identification in Communication (ASID 2010), 2010, pp. 249–252

8 Zhou, Z., Dai, W., Eggert, J., Giger, J.T., Keller, J., Rantz, M., He, Z.: 'A real-time system for in-home activity monitoring of elders'. Annual Int. Conf. of the IEEE, Engineering in Medicine and Biology Society, EMBC 2009, 2009, pp. 6115–6118

9 Saha, S., Banerjee, A., Basu, S., Konar, A., Nagar, A.K.: 'Fuzzy image matching for posture recognition in ballet dance'. IEEE Int. Conf. on Fuzzy Systems (FUZZ 2013), 2013, pp. 1–8

10 Banerjee, A., Saha, S., Basu, S., Konar, A., Janarthanan, R.: 'A novel approach to posture recognition of ballet dance'. IEEE Int. Conf. on Electronics, Computing and Communication Technologies (IEEE CONECCT 2014), 2014, pp. 1–5

11 Saha, S., Ghosh, S., Konar, A., Nagar, A.K.: 'Gesture Recognition from Indian classical dance using Kinect sensor'. Fifth Int. Conf. on Computational Intelligence, Communication Systems and Networks (CICSyN 2013), 2013, pp. 3–8

12 Saha, S., Ghosh, S., Konar, A., Janarthanan, R.: 'Identification of Odissi dance video using Kinect sensor'. Int. Conf. on Advances in Computing, Communications and Informatics (ICACCI 2013), 2013, pp. 1837–1842

13 Phung, S.L., Bouzerdoum, A. Sr., Chai, D. Sr.: 'Skin segmentation using color pixel classification: analysis and comparison', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2005, **27**, (1), pp. 148–154

14 Jones, M.J., Rehg, J.M.: 'Statistical color models with application to skin detection', *Int. J. Comput. Vis.*, 2002, **46**, (1), pp. 81–96

15 Dyer, C.R.: 'Computing the Euler number of an image from its quadtree', *Comput. Graph. Image Process.*, 1980, **13**, (3), pp. 270–276

16 Deans, S.R.: 'Hough transform from the Radon transform', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1981, **PAMI-3**, (2), pp. 185–188

17 Beylkin, G.: 'Discrete radon transform', *IEEE Trans. Acoust. Speech Signal Process.*, 1987, **35**, (2), pp. 162–172

18 Lewis, J.P.: 'Fast normalized cross-correlation', *Vis. Interface*, 1995, **10**, (1), pp. 120–123

19 Chaudhary, A., Raheja, J.L.: 'Bent fingers' angle calculation using supervised ANN to control electro-mechanical robotic hand', *Comput. Electr. Eng.*, 2013, **39**, (2), pp. 560–570

20 Theodoridis, S., Pikrakis, A., Koutroumbas, K., Cavouras, D.: 'Introduction to pattern recognition: a Matlab approach' (Academic Press, 2010)

21 Mitchell, T.M.: 'Machine learning and data mining', *Commun. ACM*, 1999, **42**, (11), pp. 30–36

22 Panigrahi, B.K., Suganthan, P.N., Das, S., Dash, S.S.: 'Neural network based gesture recognition for elderly health care using Kinect sensorj', in (Eds.): 'Swarm, Evolutionary, and Memetic Computing' (Springer, 2013), pp. 376–386

23 Pal, M., Saha, S., Konar, A.: 'A fuzzy C means clustering approach for gesture recognition in healthcare', *Knee*, **1**, p. C7

24 Dietterich, T.G.: 'An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization', *Mach. Learn.*, 2000, **40**, (2), pp. 139–157

25 Veltkamp, R.C.: 'Shape matching: similarity measures and algorithms'. Int. Conf. on Shape Modeling and Applications, SMI 2001, 2001, pp. 188–197

26 Mokhtarian, F., Abbasi, S.: 'Retrieval of similar shapes under affine transform'. Proc. of the Third Int. Conf. on Visual Information and Information Systems, 1999, pp. 566–574

27 Efrat, A., Itai, A.: 'Improvements on bottleneck matching and related problems using geometry'. Proc. of the Twelfth Annual Symp. on Computational Geometry, 1996, pp. 301–310

28 Arkin, E.M., Chew, L.P., Huttenlocher, D.P., Kedem, K., Mitchell, J.S.B.: 'An efficiently computable metric for comparing polygonal shapes', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1991, **13**, (3), pp. 209–216

29 Dietterich, T.G.: 'Approximate statistical tests for comparing supervised classification learning algorithms', *Neural Comput.*, 1998, **10**, (7), pp. 1895–1923