

e-ISSN: 2395 - 7639



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

Volume 9, Issue 7, July 2022



INTERNATIONAL **STANDARD** SERIAL NUMBER INDIA

Impact Factor: 7.580

0

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|



Volume 9, Issue 7, July 2022

DOI: 10.15680/IJMRSETM.2022.0907016

Spatio-Temporal and Support Vector Machine Based Human Action Detection

Monisa Nazir¹, Kirti Bhatia², Shalini Bhadola³, Rohini Sharma⁴

¹Student, Sat Kabir Institute of Technology and Management, Haryana, India
 ²Assistant Professor, Sat Kabir Institute of Technology and Management, Haryana, India
 ³Assistant Professor, Sat Kabir Institute of Technology and Management, Haryana, India
 ⁴Assistant Professor and corresponding Author, GPGCW, Rohtak, India

ABSTRACT: We have developed a model which can identify Human activities such as running, prodding, walking, hand clapping, boxing, and hand weaving. We have used spatial-temporal local features to extract the actions. Then we used a support vector machine (SVM) classifier to categorize different actions of a human. Our proposed method efficiently recognizes all main activities of humans.

KEYWORDS: Human Activity Recognition, Computer Vision, Spatial, Temporal, SVM

I. INTRODUCTION

Determining people's actions from videos or images is a complex procedure because of issues like contextual clutter, partial obstruction, differences in scale, outlook, illumination, and look. A multiple activity identification model is necessary for numerous applications, such as video monitoring systems, HCI, and robot for characterising human behaviour. To distinguish human behaviours, we have created a temporal and space local feature mechanism with a support vector classifier (SVM) machine learning approach. We used the KHH dataset and our method, and the outcomes are completely accurate.Examining actions from still photos or video clips is the aim of human activity recognition. This fact serves as the driving force behind person activity identification systems' quest to accurately classify input data into the relevant activity category. Human activities are divided into the following categories based on their complexity: gestures; (ii) single activities; (iii) inter-actions between humans or between humans and objects; (iv) collective actions; (v) behaviours; and (vi) events.

II. RELATED WORK

Human activity recognition is a part of a vision-based method like the identification of humans through a webcam [1-2], image quality enhancement [3], and recognition of face [4]. Active learning (AR) in HAR is a recent research area. The AL methods seek to reduce the difficulty and expense of learning. A good quantity of revealing unmarked information samples should be chosen, and the labels should be requested. It increases prediction accuracy while reducing labeling effort [5]. In contrast to other domains where active learning has gained popularity, only a small number of scholars have studied active learning in AR. To identify the most informative unlabeled data sample, Authors in [6] studied three distinct strategies and demonstrated that the annotation effort was decreased by a factor of 2. Reviews of techniques for recognising hand gestures using vision were published in [7]. These studies served to present the gesture identification as a means of communicating with devices and to give academics a rundown of developments in hand gesture identification in order to assist them pinpoint fields that still require more study. The authors of [8] concentrate on hand gesture identification technologies and also gesture taxonomies, presentations, and algorithms for recognising them. As opposed to actions and gestures, some academics are more concerned with interpreting behaviour and events. Motion estimation to identify human activities piqued the interest of other academics, and numerous reviews were published in this area. The authors of [9] make an effort to enumerate depth-based person

International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management (IJMRSETM)



| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580

Volume 9, Issue 7, July 2022

DOI: 10.15680/IJMRSETM.2022.0907016

movement analysis techniques. Person body part-based movement study, moving person tracing, and image sequencecentered person identity are three of the 3 sub topics of human movement analysis covered by [10].

III. AUTOMATIC RECOGNITION OF HUMAN ACTIONS

For content-centered video indexing, intellectual supervising, HMI, and virtual reality, human activity identification in videos are crucial. A promising use of computer vision is the automatic identification of human actions in video. Application scenarios include human-computer interaction, intelligent video surveillance, and content-based video retrieval. Although many scholars have conducted a lengthy study in this area, it is still difficult to identify human actions in films because of variations in scales, rotations, viewpoint, brightness, and occlusion in addition to geometric variances across intra-class items or motions. Extraction of features, visual representation, cross fusion, and classification are the four processes that makeup one of the most widely used frameworks for HAR. The classification procedure for recognising human activity based on spatial and temporal data is shown in Fig. 1.



Fig 1: Steps of Human Activity Recognition

SPACE-TIME TRAJECTORY

Activity is seen as a collection of space-time trajectories in trajectory-based techniques. In these methods, a person's joints and body position are represented by 2-dimensional (XY) or 3-dimensional (XYZ) points. Depending on the action being performed, a person's joint positions alter when doing that action. Space-time trajectories are used to depict these changes as a 3D or 4D XYZT representation of the action. The body's joint location is tracked by the space-time trajectories in order to discern between various actions.

SPACE-TIME FEATURES

For the purpose of recognising human action, space-time features-based approaches extract features from space-time volumes or space-time trajectories. These traits are typically local in nature and contain action-specific discriminative properties. These properties can be categorised into two groups: sparse and dense, depending on the characteristics of space-time volumes and trajectories. While feature detectors based on optical flow are thought to be dense, those based

International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management (IJMRSETM)



| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580

Volume 9, Issue 7, July 2022

DOI: 10.15680/IJMRSETM.2022.0907016

on interest points are regarded as sparse. The majority of the newly proposed algorithms are built on these interest point detectors.

KTH HUMAN ACTION DATA COLLECTION

The Royal Institute of Technology in Sweden produced the KTH dataset [11] in 2004. This dataset includes 25 actors performing six various human activities in 4 separate contexts. Thus, there are 600 video sequences total, or 2564. This dataset is also regarded as being very basic for the evaluation of human activity detection algorithms because the films were taken with a static camera and background.

PROPOSED SOLUTION

With the help of local measurements in terms of spatiotemporal interest sites, we have shown in this work that action recognition is possible (local features). Such features can adjust to the magnitude, periodicity, and direction of movement sequences, capturing local motion events in videos and producing representations of video that are stable with regard to associated transformations.

Furthermore, local characteristics and SVM have been incorporated in a reliable classification method for spatial recognition. Corresponding to that, in this case, we investigate the integration of real time-space characteristics with SVM and use the subsequent methodology to recognise human actions. We provide a new video database for evaluation purposes and report findings from identifying six types of human activities carried out by 25 unique humans in various settings.

We build the scale-space representation of a picture sequence using a Gaussian convolution kernel to identify local features. The space and time variables of the related Gaussian kernel then determine the space and time neighborhood of features in space and time. By automatically choosing scaling parameters, the size of features can be adjusted to fit the Spatio-temporal extent of underlying image structures. Additionally, the features' shapes can be adjusted to match the local pattern's velocity, making them stable in the face of varying degrees of camera motion. The velocity and spatial representation of activities in consecutive frames are described by the spatiotemporal neighbourhoods of local characteristics. We compute spatio-temporal jets at the center of each feature using normalised derivatives to encode this data. Prior to computing descriptor 1, we also warp the neighbours of features using predicted velocity values to ensure invariance with regard to related camera movements. A lexicon of basic events is produced by K-MCg of the descriptors 1 in the training set. A feature histogram is defined by the quantity of features with labels hi in a specific order ($h_1, ..., h_n$). When identifying motion in image sequences, we employ such histograms as one possible expression.

SUPPORT VECTOR MACHINES (SVM) BASED CLASSIFICATION

Modern big margin classifiers, such as the SVM, have lately grown in favor of the field of visual pattern recognition. Take into consideration the issue of categorising the training data set into two classes, where xi _N is a feature vector and yi 1, +1 its class label. The best hyperplane is the one that maximises the margin if we suppose that the two classes may be divided by a hyperplane of the form w x + b = 0 in some space H and that we are unaware of the distribution of the data. By employing Lagrange multipliers i(i = 1,..., m), it is possible to find the ideal values for w and b by resolving a restricted optimization technique.

IV. EXPERIMENTATIONS AND RESULTS

The experimental setup, training procedure, and experimental outcomes for the suggested technique are covered in this section. On well-known action datasets like KTH, the suggested approach is tested. In relation to local characteristics, SVM classification has been integrated with motion descriptors (LF). Our concept was put into practice using



| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580

| Volume 9, Issue 7, July 2022 |

DOI: 10.15680/IJMRSETM.2022.0907016

MATLAB 2014 software. The KTH dataset's photos were then used to examine our results. We have selected a video frame from KTH database.Fig 2 shows recognition of walking activity.Fig 3 shows recognition of cycling by a human.Fig 4 shows the surfing activity of a human.Fig 5 shows recognition of boxing activity of humans.Fig 6 shows hand clapping by humans and Fig 7 shows hand waving by a human.Fig 8 and Fig 9 recognize jogging and running activities of human.

B R2014a					
PLOTS APPS EDITOR PUBLISH	VIEW			1 1 3 c 5 (Search Documentation
Tind Files Insert 🛃 f 🌾 🖌 🗘	akpoints Run Run ar	Run Sec Figure 1			
File Tools View Playback Help	👻 💌 Advan	File Edit View Insert Tools	Desktop Window	Help	×
	n Code 🕨		9 4 14 - 3		
	ng fram Help Dia				9 9 I
	ing fram ing fram ing fram ing fram ing fram	Walking			н н
	ng frame no.32 ng frame no.33 ng frame no.34 ng frame no.35	8 - 9 - 10 -	7 7 7		а а
	.ng frame no.36 .ng frame no.37 .ng frame no.38 .ng frame no.39	11 - 12 - 13 - 14 -		n - 1	a 3
Value Min Max Proces Value Min Max Proces 50x40 double 25.5848<6.6600	ing frame no.40 sing frame no.41 sing frame no.42 sing frame no.43 sing frame no.44	10 - 16 - 17 18 19 - 19 - 20 -			. т
50x40 double 32.4654 6.8500 E Proces	sing frame no.45	20 - 21 - mov(k).cdata = read(1	, k);)

Fig 2: Recognition of Walking Activity

AB R2014a	C. R. W. P	Real Property lies			100000000000000000000000000000000000000										
IE	PLOTS	APPS EDITO	R PUBLISH	VIEW		Figure 1									23
Movie I File Too	Player ols View Playb	ack Help		eakpoints Run	Run Section Run and Advance	File Edit	View Inser	rt Tools	Desktop	o Wind	low He	elp 📰 🔳 🕻]		Ľ
H 44 4		ਮ ⊡ ¢ ⇒ 0 ਯੋਗ ਹੈ ਹੈ	/ [] 100%	on Code >	RUN		ti ki	h	ħ		Ba	n i	o Ba	ħi.	
			U	ing frame ing fram ing fram ing fram ing fram ing fram	Help Dialog	-	×		ħ.	h	la	n 1	i 14	Re .	
			14 AN	ing fram ing fram ing frame f ing frame r ing frame r	OK 10.32 8 - 10.33 9 -		. Ka		Ka	h	Ba	n I	a h	Ba	
		$\langle \rangle$	MAGENER	ing frame r ing frame r ing frame r ing frame r	10.35 10 - 10.36 11 - 10.37 12 - 10.38 13 -		ta ki	h	Ro.	h	lu -	6 B	a ki	la	
Stopped ce	Value	RGB:256x256	(100% (15 fps) 1 / 50 Proces Proces Proces	ing frame r sing frame r sing frame r sing frame r sing frame r	10.39 14 - 10.40 15 - 10.41 16 - 10.42 17 10.43 18		ts Ko	h	B a	h	ia.	ta i	a ka	Ba	

Fig 3: Recognition of Cycling Activity of Human



| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 7, July 2022 |

| DOI: 10.15680/IJMRSETM.2022.0907016 |

MATLAB R2014a	and the second se		
HOME PLOTS APPS EDITOR PL	IBLISH VIEW	Higure 1	
Movie Player	Run Section	File Edit View Insert Tools Desktop V	Window Help ∞ ▼ 🕃 🗈 🔲
	Breakpoints Run Run and Advance		
	ognition Code > and Window Doessing fram Doessing fram Doe		· · · · · · ·
Alasta and a second	beessing frame no.35 9 - 1 beessing frame no.35 10 - 1 beessing frame no.36 11 - 1 beessing frame no.37 12 - p	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	
Stopped [RGB:256x256 100% (15 fps) 1 / 50	cessing frame no.38 13 - n ocessing frame no.39 14 - v		
ognize.m (Script)	Processing frame no.40 15 - v Processing frame no.41 16 - m	1	
pitspace O	Processing frame no.42	L	

Fig 4: Recognition of Surfing Activity of a Human.

📣 MATLAB R2014a				
Movie Player	JSH VIEW	Figure 1	Contraction of the local division of the loc	
File Tools View Playback Help 🏻 🛪		File Edit View Insert	Tools Desktop Window Help	¥ ا
🗋 🚔 🖆 🖞 🔞 🕕 💁 🔍 🤍 💭 🔝 100%	Breakpoints Run Run and Advance	12 😋 🖻 🖗 🔗 🤗	° (*) (*) (*) (*) (*) (*) (*) (*) (*) (*)	
$ \mathbf{H} \triangleleft \mathbf{H} = \mathbf{F} \mathbf{F} \triangleleft \mathbf{H} \mathbf{G} \mathbf{G} \Rightarrow$	▼ ▼ Advance			
	E BREAKPOINTS RUN			
4	nand Window 💿 📝 Editor			
	cocessing frame no.25			
	cocessing frame no.26 1	8 E	. n. z. n. z. n	т т т т
	🛃 Help Dialog			
	Boxing			
	ОК			
	cocessing frame no.36 11 -			
Stopped RGB:120x160 100% (25 fps) 1/408	cocessing frame no.37 - 12 - 13 -			

Fig 5: Recognition of Boxing Activity of Human.

LAB R2014	a	A Manual		_	-		-		the state of the s													- 0
оме	PLOTS	APPS	EDITOR	PUBLISH	V	EW				🚺 Figu	ire 1					-	-	-				23
🛃 Movie	Player	Seattle .					R	Run Se	ction	File	Edit V	liew	Insert	Tools	Desktop	Wind	low I	Help				۲
File To	ols View Playb	ack Help		r	Breakpoints	Run	Run and	Advan	ce Run i	20				R 🔊	94.		3					
10	🗃 🗄 🖷 🔞	1 🗟 🕰	Q 👋 🖾 100	1%		•	Advance	DUN	Tim													
14 44 4	I = > I> >>	H D C	4		nition Code	۶.		RUN	_													
					Window	-	e	Edito	r - F:\MATL			т		T.		T.				T.		
					SSIIIg II	dalle in	0.21	+4	Raconnize.	9												
					i i i i i i i i i i i i i i i i i i i	alog				1		Ŧ		Ŧ		T.		a.		a .		
						Linud	Classics		-													
		-				Hand	Clapping															
		T					ОК]				т		т		т		т		Ŧ		
		Ĩ						8 -	deleta													
	Letter 1				ssing fr	ame no	0.33	9 -	[filer	8												
					ssing fr	ame n	o.35	10 -	I = V:	a)		т.		т	ч.	Ŧ	*			т		
					ssing fr	ame n	0.36	11 -	pause													
					asing fr	ame no	0.37	13 -	nFrame	8												
Stopped		RGB:1	20x160 100% (25	5 fps) 1/312	ssing fr	ame n	0.39	14 -	vidHei	i i		я.		T.			Ψ.		7	т	,	
lize.m (Scri	pt)		~	Proc	essing fr	ame n	0.40	15 -	vidWid	8												
pace			0	Proc	essing fr	ame n	0.41	16 -	mov (1:													
	Value	Mir	n Max	Proc	essing fr	ame no	0.42	18	3					_								

Fig 6: Recognition of Hand Clapping Activity



| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

Volume 9, Issue 7, July 2022

| DOI: 10.15680/IJMRSETM.2022.0907016 |

A MATLAB R20	14a	- A MARKET	-	-						Figure	1									
HOME	PLOTS	APPS	EDITOR	PUBLISH	VI					File Ed	dit View	Insert	Tools	Deskto	p Wind	dow	Help			×
Movie Play	yer			23			NE.	~		กล			 € 	1	1 - 1	al		-		
File Tools	View Playba	ck Help		× ×	2	2		Run S	ection			~ ~	* 57	O R	* • 1	Geo L				
	日 增 🔯	🕕 💁 🔍 Q	🆑 🖾 100%	-	Breakpoints •	Run	Run and Advance	Advar	nce F											
19 99 91	= 🕨 IF FF	ы ⊕ ⇔ ≑	2	E	BREAKPOINTS			RUN									4		T	
4				ecog	nition Code	•														
				mand	Window	_	e	Edit	or - F:\M	6										
				rd 🛛	Help Dialo	g	100		× 1	8										
				rd l	~				P				т		7		т		T.	
				co					- 11											
				co	\triangleleft	and Wav	ing		2	1										
		-		rd		_			- 10						14				Ŧ	
				co			OK													
	A STATES			rd			20	7 -	clo											
	Carlos Carlos			roce	essing fr	ame no	.33	8 -	del	ē										
				roce	ssing fr	ame no	.34	9 -	[fi	6	ч		т				1		4	
				coce	essing fr	ame no	.35	10 -	I -											
				roce	ssing fr	ame no	.36	11 -	imp	4										
				roce	essing fr	ame no	.37	12 -	pau	8			-		-		-		iii	
Stopped		RGB:120x	160 100% (25 fps)	1/581 roce	ssing fr	ame no	.38	13 -	nFr	9										
				Froce	essing fr	ame no	.39	15 -	Vid											
Recognize.m (So	cript)		^	Proce	essing fr	ame no	.40	16 -	mov											
Workspace			۲	Proce	esing fr	ame no	42	17	1			_					-			

Fig 7: Recognition of Hand Weaving Activity

B R2014a	100					J Figure 1									L	
PLOTS APPS EDITOR PUBLISH	VIE	W				File Edit	View	Inser	t Tools	Desktop	Windo	w He	lp	_		¥د ا
Movie Player					2	163		2 🕀	. e. 80	9 .	4 • 0					
File Tools View Playback Help 🔉			EY	Run Section	0					-	- 1					
🎦 🚍 🖆 🛱 🛱 🔞 🗿 💁 🔍 🔍 🖏 100%	ikpoints •	Run	Run and Advance	Advance	Ru											
K 44 4 ■ ► B + B 1Ω 40 ₹	KPOINTS	_		RUN												1 mg
	h Code 🕨															
a	dow	_	۲	Editor - F:	MA											
5	Hel	p Dialog	1		~											
5																
		🛃 Joj	gging													
			_	_			104	-				-	-	142		
			0	IK												
9	ng fra	me no	. 32	7 - cl	Lc											
	ng fra	me no	.33	8 – de	ele											9.44
	ng fra	me no	.34	9 - [1	Eil(1	*			<u>_</u>		3			
i la	ng fra	me no	.35	10 - I	= 1											
j.	ng fra	me no	.36	11 - in	npl											
g	ng fra	me no	.37	12 - pa	aus											10.02
Stopped RGB:120x160 100% (25 fps) 1/420	ng fra	me no	.38	13 - nH	Frai											1.
d	ing fra	me no	.39	14 - vi	LdH											
a.m (Script) ^ Process	sing fra	me no	.40	15 - vi	LdW.											
Procesi	sing fra	me no	. 41	16 - mc) va											
e 🐨 Process	aing fra	me no	42	17	4			-					_			

Fig 8: Recognition of Jogging Activity

LAB R2014a		
	Plinise VEW	
Movie Player File Tools View Playback Help 그	Image: Section for the section	×
	REMPOINTS RUN Ition Code > Window STURG Frame, no. 4	
	est Help Dialog	
	ssi ssi ssi ssing scame norse	a, W
	sing frame no.33 8 - delete('Fi sing frame no.34 9 - [filename sing frame no.35 10 - I = Videol sing frame no.36 11 - implay([pa sing frame no.36 12 - pause(3)]	н т.
Stopped RGB:120x160 100% (25 fps) maicmxmm ize.m (Script) ^ ace ·	1//325 psing frame no.37 1/325 sing frame no.38 14 vidHeight Processing frame no.40 15 Processing frame no.41 16 16 mov(1nFrame no.41)	

Fig 9: Recognition of Running

LIMBSETM

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580

Volume 9, Issue 7, July 2022

DOI: 10.15680/IJMRSETM.2022.0907016

V. CONCLUSION

Despite differences in clothes, lighting, and personal movement patterns, the localization of descriptors enables matching of similar events. But, some of the matched characteristics belong to various parts of (different) actions that are challenging to identify based just on local information because of the local nature of features and their corresponding jet characteristics. Therefore, it is clear that taking into account the spatial and temporal coherence of local characteristics will strengthen our strategy. We have shown the representation and recognition of motion patterns, such as human motions, using local spatiotemporal characteristics. We developed a novel motion recognition methodology that outperforms competing approaches in terms of recognition performance by merging local information with SVM. The benefits of representing gesture forms in regard of common attributs is that they can withstand changes in the pattern's scale, periodicity, and speed.

REFERENCES

[1] Nakul Nalwa, Shalini Bhadola, Kirti Bhatia, Rohini Sharma, Development of Vision Based People CountingFramework using Webcam, International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET), Volume 5, Issue 7, July 2022, pp. 1665-1671.

[2] Nakul Nalwa, Shalini Bhadola, Kirti Bhatia, Rohini Sharma, A Detailed Study on People TrackingMethodologies in Different Scenarios, International Journal of Innovative Research in Computer and Communication Engineering, Volume 10, Issue 6, June 2022, pp. 5907-5913.

[3] Mahesh Kumar Attri, Kirti Bhatia, Shalini Bhadola,Rohini Sharma, DEVELOPMENT OF A FOURIER TRANSFORMATION BASED MODELFOR IMAGE SMOOTHING, International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:06/June-2022, pp. 4204-4210.

[4] Sanju, Kirti Bhatia, Rohini Sharma, An analytical survey on face recognition systems, International Journal of Industrial Electronics and Electrical Engineering, Volume-6, Issue-3, Mar.-2018, pp. 61-68.

[5] Hossain, H. S., Khan, M. A. A. H. and Roy, N. Active learning enabled activity recognition. Pervasive and Mobile Computing, 2016.

[6] Alemdar, H., van Kasteren, T. and Ersoy, C.Active learning with uncertainty sampling for large scale activity recognition in smart homes. Journal of Ambient Intelligence and Smart Environments, 9, pp. 209–223, 2017.

[7] Cheng H, Yang L, Liu Z, Survey on 3d hand gesture recognition. IEEE Trans Circ Syst Video Techn2016, 26(9):1659.

[8] Rautaray SS, Agrawal A, Vision based hand gesture recognition for human computer interaction: A survey. ArtifIntell Rev, 2015, 43(1):1.

[9] Moeslund TB, Hilton A, Krüger V (2006) A survey of advances in vision-based human motion capture and analysis. Comput Vis Image Understand 104 (2-3):90.

[10] Kratz, L. and Nishino, K. (2012). Tracking pedestrians using local spatio-temporal motion patterns in extremely crowded scenes. IEEE transactions on pattern analysis and machine intelligence, 34 (5), pp. 987–1002.

[11] Liu,J., Yan,J., Tong,M., and Liu,Y.(2010). A Bayesian framework for 3D human motion tracking from monocular image, in IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX:IEEE, pp. 1398–1401.









INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT



+91 99405 72462



www.ijmrsetm.com