

Review

Biologically-Inspired Computational Neural Mechanism for Human Action/activity Recognition: A Review

Bardia Yousefi ^{1,2}  and Chu Kiong Loo ^{2,*} 

¹ Computer Vision and Systems Laboratory (CVSL), Department of Electrical and Computer Engineering, Laval University, Quebec City, QC G1V 0A6, Canada; Bardia.Yousefi.1@ulaval.ca

² Advanced Robotic Laboratory, Department of Artificial Intelligence, University of Malaya, Kuala Lumpur 50603, Malaysia

* Correspondence: ckloo.um@um.edu.my; Tel.: +1-267-669-9954

Received: 30 September 2019; Accepted: 8 October 2019; Published: 15 October 2019



Abstract: Theoretical neuroscience investigation shows valuable information on the mechanism for recognizing the biological movements in the mammalian visual system. This involves many different fields of researches such as psychological, neurophysiology, neuro-psychological, computer vision, and artificial intelligence (AI). The research on these areas provided massive information and plausible computational models. Here, a review on this subject is presented. This paper describes different perspective to look at this task including action perception, computational and knowledge based modeling, psychological, and neuroscience approaches.

Keywords: human visual system; recognition of biological movement; human action recognition; biologically inspired model; ventral and dorsal streams interaction

1. Introduction

Analysis of biological motion recognition is categorized in different research fields, such as neurophysiology, neuropsychological, artificial intelligence (AI), and computer vision. The present article mainly investigates the AI perceptive of this task concerning neurophysiology, and neuropsychological evidence. This considers a subdivision of a computer science field dealing with machine intelligent behavior and learning. The primary experimental methods on this field have been classified into two main divisions:

- (1) **symbolic methods**, which follow classic approaches, are similar to expert systems and termed as connectionism approaches;
- (2) **scruffy methods** concentrate on the intelligence evolution or following artificial neural networks.

Both directions undergo rigorous restrictions. The untimely objectives, such as human behavior reproduction and simulation, are entirely overlooked. Machine and man have opposite abilities: man can estimate, infer, and recognize in parallel, whereas machine can sequentially perform quick computations. Biologically inspired models are similar to behavior-based AI, and have attracted attention of scientific communities. These models are more focused on the performance than on the internal processing of the machine. Many projects have intensive collection of the mentioned facts, but none can create a machine that takes direct advantage of this information. Modern machines can learn on the basis of statistics and follow few determined objectives. One objective is to deal with information extraction from large datasets, unsupervised learning, pattern recognition, and calculation based on statistical analysis statements. Such machines are practically used for actual problems, such as speech, image, and object recognition. Biological movement and its recognition comprise a multifield

research that follows many biological principles and engineering approaches, and are based on the human (or mammalian) visual system. In the present study, computational model for recognition of biological movement is further reviewed in terms of mechanisms and models proposed for this task.

2. Motivation

Human action recognition in monocular video is an important subject regarding video applications, such as in human computer interaction, video search, and many other relevant tasks. Biologically-inspired recognition models provide more insight regarding the performance of such approaches compared to the mammalian visual system following neurophysiological, and psychophysical evidence. Understanding about the mammalian visual system provides an opportunity to improve the artificial intelligence (AI) models to be stronger with higher similar capability to human. This paper reviews different aspects of the biologically-inspired mechanisms for recognition of human action into the following overall categories:

- Perception of the motion
- Knowledge based modeling of the human action
- Psychology and neurophysiology of the motion

To better follow the different aspects of biological human action, these subjects are divided into several smaller subsections (there are other similar categorizations that exist i.e.,[1]).

Figure 1 represents a categorization of human action recognition based on neuro-physiological evidence and computer based modeling.

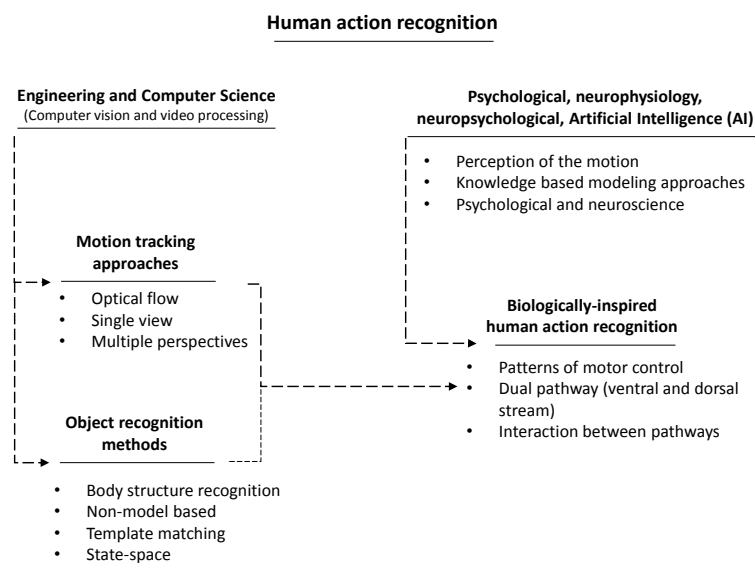


Figure 1. A categorization of human action recognition methods in different methodological perspectives.

3. Analysis of Biological Movement

3.1. Motion Patterns

Human action recognition is summarized by automatically determining the type of human action (human as a moving object in video sequences (image frames)). Motion recognition has been studied in neurophysiological, psychophysical, and imaging experiments. Marey and Muybridge carried out the initial studies on human movement in the 1850s; they photographed moving subjects and presented on its locomotion [2]. One of the earliest studies on visual perception and introduction of actual movement was conducted by Rubin (1927) and Duncker (1929) [3,4]. Johansson et al. (1973) [5] presented the initial part of a study on moving object characteristic and movement perception (similar to [6]). In general, the methods for motion patterns can be categorized into two different paths:

1. One technique uses global feature extraction from video streams to allocate a particular label to the whole video. This technique clearly needs an unchanged observer within the video, and the environments where actions are occurring should be considered [7].
2. The second technique considers local features in each frame and label for distinct action. Afterward, sequences can be attained through simple voting for global labeling. Temporal analysis for obtaining the features in each frame and classification is based on the observation in temporal window.

Both approaches showed significant outcomes [8–10]. Learning is also fundamental in recognizing 3D stationary human motion [11]. Human action recognition using video frames categorizes as an object recognition problem. Such recognition is supposed to handle object variations (e.g., style and size), and the human brain can excellently categorize human objects in different classes of action (recent bio-inspired methods in computational neuroscience [12,13]). In the primary visual cortex (V1), the image procedure is particularly sensitive on bar-like structures (Gabor-like methods and showed in Figure 2). V1 responses are combined by extrastriate visual areas and passed to the inferotemporal cortex (IT) for recognition [14].

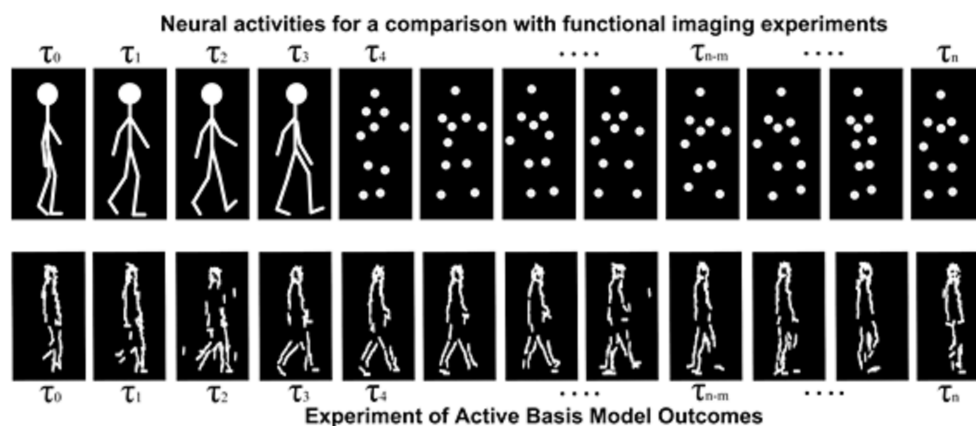


Figure 2. A illustrative similarity comparison of MLD and ABM [15] from the perception of biological movements. These basis have considerable consistency to the point light technique (MLD), which presents static pictures. ABM also has a good representation of biological movements in form pathway. Picture is adapted from [16].

3.2. Kinetic–Geometric Model

Kinetic–geometric model presents the analysis of visual vector and basically expansion in mechanical form when biological motion perception and its patterns are combined. A classic moving light display (MLD) provides an excellent impetus for human motion perception in neuroscience analysis and study [5,17]. Recognition of human walker’s gender has been performed without knowledge cues by point light sources established on important joints of the human body, unlike versions of statistical experiments that have sufficient accuracy in this task. Changes in the speed of walking and degree of arm-swing especially in higher speeds are associated with females and upper body joints in lights analysis which are better in finding the accuracy for gender recognition [18]. Marr et al. (1978) showed the problem of computational process in human visual system and information obtained via retinal images; 3D shapes are considered for problem presentation by introducing some notes as follows: three criteria are introduced for shape recognition in judging, three aspects of design representation are considered (e.g., coordinate, primary shape unit information, and information organization), shape description (e.g., coordination of the object center, size variations, modular organization, view transferring mechanisms, and identification of natural axes), and constraints for conservation recognition applying further information from the image. Perrett et al. (1985) reported on the temporal cortex of macaque monkey; they found that most of the cells in the

brain region are sensitive to the type of movement and respond to specific body movements [19] (models showed similar behavior applying different receptive field showed in Figure 3). Two cell types introduced are sensitive to the rotation and view of body movements and the response of majority of cells in these areas of temporo-parieto-occipital and PGA of temporal cortex has been considered for providing descriptions of view-centered and view-independent responses among the mentioned cells [20]. Goddard (1989) used connectionist techniques, along with spatial and temporal feature incorporation through diffused MLD data, and represented the walking recognition in 400 ms MLDs [21]. This integration occurs in the low-level features of shape and motion by this target to make high-level features. Low-level features include sequential trajectories points, and they are grouped in line segmentation with others to obtain proper lower and upper body limb forms (or even correlated features [22]). Figure 2 represents the similarity between MLD and active basis model (ABM) [15,16]. An unsupervised approach performs synthetic biological movement recognition [23] and shows great potential for use in the mechanism of biological movements and the importance on geometric model implying synthetic data.

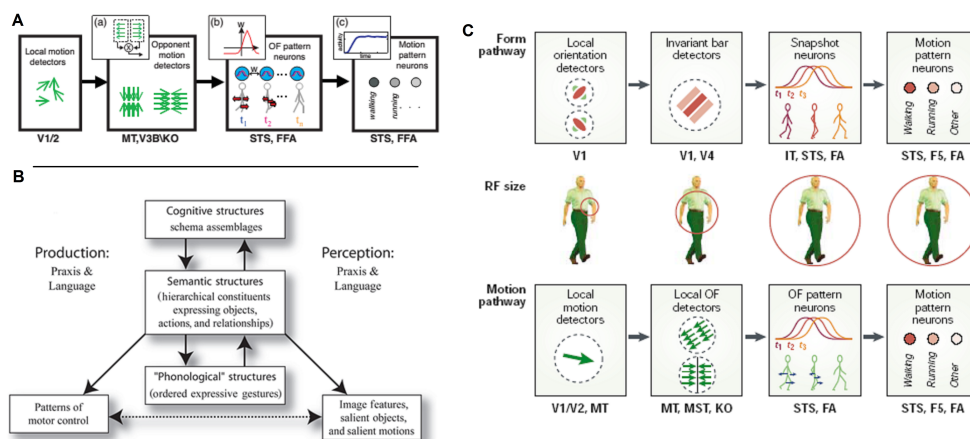


Figure 3. (A) Schema of the model shown and symbols are shown following the brain areas and their functionality: MT: middle temporal area; V1: primary visual cortex; FFA: fusiform face area; STS: superior temporal sulcus; KO: kinetic occipital area. These areas and their functionalities are considered in their timing t_1, t_2, \dots, t_n for the input data frames and their encoded information gathered by radial base function and optical flow. In addition, (a) reveals the opponent motion detector; (b) shows the lateral coupling in complex optical flows; (c) response of the motion pattern detector. [24]; (B) The language production and perception performance review is shown in the figure [25]; (C) illustrate the model presented in [26] with concerns of receptive field as well.

3.3. What and Where Pathways

Shape and form pathways are hierarchically joined to detect the three levels of complexity, i.e., component, segment, and assembly levels that signify temporal series on procedures [21]. Goddard followed the biologically inspired human action recognition in determining the complex structured motion by using MLDs. He analyzed major computational problems, such as time-varying representation, visual stimuli integration, gestalt formation, contextual formation, and particular spatial location focusing on process and its representation. Moreover, he showed the process of “what” and “where” in the visual system tightly coupled in a synergistic manner [21,27] (we will discuss this in next sections). One of the famous biologically inspired model [26,28] has proposed two independent pathways, which model the dorsal and ventral processing streams in the mammalian visual system. Figure 2 shows a representation of active basis model (ABM) and its similarity with MLD. The form pathway in ventral stream, uses Gabor filter like function to obtain the shape and form information

and as a good representation of simple and complex cells (Figure 3A,C). Motion pathway in the computational mechanism usually uses for optical flow to extract motion information.

4. Perception of The Motion

4.1. Perception And Actions

Action perception for the purpose of recognition is initially (in separate paths) presented by Goodale and Milner (1992), whom were proposed a separation of the perception and actions for recognition and identification on the ventral processing stream concerning the object recognition task. This separation allows the observer to move the hand for picking up of object and considers the projection perceptual information for object identification from striate and IT; furthermore, the posterior parietal region of the striate cortex has dorsal stream projection and needs visual sensorimotor transformations [29].

After that, Cédras and Shah (1995) reviewed motion development from the recognition aspect and emphasized on two main stages, which were presented for motion-based recognition by organizing it into motion models and matching unknown input with the constructed model. Several recognition, such as cyclic motion detection and recognition, hand gesture interpretation, tacking, and human motion recognition, were reported [30]. Perkins (1995) presented real-time animation, along with rhythmic and stochastic noise, for conveying only the texture of motion; this research also avoided the computational dynamic and constraint solvers [31]. He showed that each action has an internal rhythm and transition among movements, and realized a real-time animation. The detection of cyclic motion in frequency domain techniques of the magnitude information of Fourier transform and autocorrelation is represented as a curvature (1D signal) in function of time at 2D trajectories. Such detection is tested by synthetic and actual data of a walking person [32]. Such techniques brought this field in the motion capture direction.

4.2. Motion Patterns for Perception Of Action

4.2.1. Spatiotemporal Filter

Hubel and Wiesel (1968) suggested the occurrence of object recognition from simple to complex cells, which continued by Riesenhuber and Poggio (1999) [33,34]. But before that, Gallese et al. (1996) analyzed the electrical activity in the brain of macaque monkeys from 532 rostral parts of six inferior neuron areas and found out that the mirror neurons form a system aimed at matching observation, motor action execution in action recognition [35]. Quantitative modeling was conducted on biological feasibility for high-level recognition of the object (the model is based on MAX-like operation) [33] which showed more biological plausible than 3D object recognition in in man, monkey, and machine presented by Tarr & Bulthoff (1998) [36]. This method was expanded to a biologically inspired model approach for human arm movements and human action high-level abstraction by using the hierarchy of artificial neural networks and demonstrated the abstraction occurs in the visuo-motor control area of the brain and detected 37 degrees of freedom and biomechanical simulation with humanoids [37]. Having such approaches strengthened the idea of Gabor-like filter (or in general using spatiotemporal filters for form pathway). Figure 2 is a great representation of complex-cells implemented by ABM in form pathway [38]. The combination of spatiotemporal filter and hidden Markov model (HMM) technique was presented for MLD identification and provides decision based on the spatiotemporal sequence of the observed object features, and relatively little spatial information is caused by the segmentation of MLD image sequences, along with object identification; such information is highly temporal and is accessed by the HMM system, a major high classification rate [39] (was similar to [40]). An investigation on the spatiotemporal generalization of biological movement perception revealed the response of motion stimuli and interpolated this generalization among natural biological motion patterns [41]. A linear combination of spatiotemporal patterns is estimated using natural movement

patterns. The weight of prototypes in the morphs and the continuous and smooth variations in category probabilities are observed in this approach. A generalization exists within the motion patterns classes in the visual system [42], which also used for activity recognition [43]. However, these methods could not find more popularity than bio-inspired model-based spatiotemporal interesting points for human action recognition proposed by Cai et al. (2014) due to following the dual pathways and comparison between such points with the spatio-temporal interesting points (STIP) framework [44].

4.2.2. 3D Structural Method

The problem in most conventional methods might lie in the difficulty of creating computer-generated characters that display real-time, engaging interaction, and realistic motion, as well as the process of action perception in a common representational structure. Some methods tried to solve this issue by analyzing the action in 3D structure (such as [36]). The interpretation of human motion is divided into three following tasks detecting human body parts [45], tracking body using single/multiple cameras, and recognizing human activities through frame sequences, instead of simple translational model [46]. That involved low-level segmentation of human body parts involves joints and projection of the 3D structure of the human body in 2D representation [1,45,47]. Also this method could be more space-time dependent (like [28]) or geometry based movements approaches (such as [48]). A linear non-separable effects on 2D and 3D shape targets for visual search [49] and dynamic 3D recognition [50] were found to be decomposed under the law of perceptual organization. The understanding of action observation (in the observers) in another individual has a similar neural code used to produce the same actions. Evidence of this hypothesis includes brain image studies and examination of the functional segregation light of the perceptual mechanisms subtending visual recognition and the same mechanism used for action [51]. Same methodology with slight extension used for proto-object model based on the saliency map, which involved depth information using the 3D eye tracking datasets presented by Hu et al. (2016) [52], which was followed the evidence that supports the influence of eye movements on shape processing in the human visual system [6].

4.2.3. Motion Capture

Optical capture with marker: Applying MLD previously summarized and categorized a traditional optical capture technique using marker. Attaching markers to the actor's body to have the 3D anatomic human body is used for skeleton visibility assessment (fitting), optical markers, motion capture, or mapping motion onto skeleton [53–57]. More advanced cameras helped on performing such analyses (i.e., Pan-Tilt Camera Tracking [54]). Some of the approaches here were not necessarily used for recognition of human action but applied for augmented reality or graphic applications [58,59], despite of employing deep network that can be justified as a biological-inspired model [60]. However, some of trajectory labelling applying permutation learning (using deep learning) could be more relevant [57].

Motion capture markerless: Current developments on the computer vision and imagery systems mitigate the necessity of using markers to assess the motion capture. Many of the summarized research works here involved markerless optical capture [61–64]. Giese and Poggio (2000) presented morphable models by linear combination of prototypical views to recognize biological movements and image synthesis for stationary 3D objects and its involvement in complex motion patterns. The linear combination of prototypical image sequences is used to recognize action patterns (even complex movements). The mentioned new approach can be used to analyze and synthesize the biological movement, which involves the actual and simulated video data and various patterns (which has local properties of the linear vector space) of locomotion [65]. Moeslund and Granum (2001) conducted a comprehensive survey on motion capture involving computer vision. This survey targeted an overview on the taxonomy of system functionalities and summarized it into four processes (initialization, tracking, estimation of agent pose, and recognition); each of these processes is divided into its subprocesses and various categories [66]. Moreover, Grèzes et al. (2001) analyzed human

perception in biological motion. Their study considers key role for action interpretation, identification, and predication. The main hypothesis of their approach lies under neural network specifications and its verifications through fMRI for 10 healthy volunteers. Seven types of visual motion displays are used: random dot cube, drifting random dots, random dot cube with masking elements, upright point-light walker display with masking elements, inverted point-light walker display with masking elements, upright point-light walker, and inverted point-light walker. In this approach, the hemodynamic responses of both rigid and non-rigid biological motions are connected (rigid motion responses are localized posteriorly to the rigid responses). The left intraparietal cortex is involved in non-rigid biological movement perception and associated with the posterior superior temporal sulcus (STS) and left anterior portion of Intraparietal sulcus-IPS responses. Regions, such as LOS/KO, MT/V5, and the posterior STS, are included in these activations [67]. An examination on the visual perception effects used point-light display for the arm movements of two actors for knocking and drinking movements. These actions were performed in 10 various effects. The point-light animation influenced by the phase-scrambled and upside-down versions of actions was shown to the actors for classification. The experimental results indicated that perception affects the corresponding action kinematics and movement of the phase related to the different limb segmentation [68]. In general, this part considered to be related to shape assessment pathway, whereas even with motion information can be tracked [69]. Kinect and improved Vicon cameras showed considerable contribution on markerless motion capture like [64,70]. Figure 2 represents another markerless shape capture using ABM compared to MLD. Despite similarity of ABM (or similar approach for form pathway) and MLD, motion information in the biological model extracted from optical flow (which discussed in further sections).

4.3. Computational Models

Computational approaches for biological movement perception presented in the form of the computer–human interface algorithm [71], methods which involve kinetics of human body [72] or conducted visual system analysis on neural mechanisms, anatomical and functionality into two forms, and motion distinct pathways [73]. All of these approaches are considered to be biologically inspired approaches particularly last groups. Grossman and Blake (2002) conducted visual system analysis on neural mechanisms, anatomical and functionality into two forms, and motion distinct pathways. The analysis of point-light animation involves the mutual perception of form and motion within the act of biological movement. Their study referred to a previous work regarding the activation of posterior superior temporal sulcus (STSp) and presented a new finding for the activation of fusiform (FFA) and occipital within the biological movement and generation of neural signals, which can differentiate a biological motion from a non-biological one. LOC and EBA involved in human form perception were also presented. The neural in the form and motion pathways causes the biological motion perception [73]. Jastorff et al. (2002) proposed an approach to investigate the recognition process during neural mechanism and whether the brain can learn a completely new complex pattern of action. They generated a new artificial–biological motion using the linear combination of time–space prototypical trajectories recorded through motion capturing (similar to [1,28,45,47]). This method provides a significant improvement in discrimination for all the stimuli and showing that the human brain can learn entire novel action patterns [74]. Another salient foreground trajectory extraction method used saliency detection and low-rank matrix recovery for learning the discriminative features is proposed by Yi et al. (2017) [75].

4.3.1. Form Pathway

The ability to recognize moving human figure using moving point light is considered a biological motion perception and an evidence for processing form information on body shape and local motion signals to such a vivid perception. Most of the computational mechanisms considered an independent pathway for form information [1,24,28,45,47,73,76]. This section summarized approaches focused more on form information perception. Beintema and Lappe (2002) analyzed the perception of form

pattern of human action through moving light points. The biological motion stimulus follows limited time perception of human motion without the local image motion, and the direction of the walker and walking figure coherence [77] which also followed by Kilner et al. (2003) having hypothesizes that the overlap between observation and execution causes inconsistent performance [78]. The complex shape orientation and local geometric attribute perceptual integration was also presented for global representations through two-part shape adjustment and focused more on shape analysis of visual system [79] and modified for only global form representation in [80]. A transfer perception for shape information (object shape) is presented in [81] and argued that the shared components of objects cause a high level of recognition in the objects, but the component transfer between the objects has limitation. One of the impressive evidence on form perception is presented in a research on motion blind patient (LM patient) who was suspected human homolog of V5/MT concerning the moving stimuli which the patient did not report spatial disposition of the actor (using moving lights) and ability for figural segregation on the movement basis cue and interpretation of the movements (moving parts) independently [82]. An analysis of body postures in different viewpoints and human identification using four experiments concluded that people who can identify the actions are basic-level objects and that an abstraction occurs in the visual system [83] (like point lights [84]), whereas the low presentation complexity and speed of pattern categorization and cognitive processing for peripheral vision for low-level functions are also shown to be more relevant to form perception [85]. This also has a relationship between biological motion and control unpredicted stimuli which involves shape perception, motion neural subtractions, and motor imagery found in the lingual gyrus at the cuneus border [86,87]. A study about single cells, neuroimaging data function, and field potential records shows the visual mechanism in STS in primates and humans; it also simplifies biological motion display by using point-light markers on the limbs of walkers [87–89].

4.3.2. Motion Pathway

The motion perception has been investigated in the aspect of psychology by Giese (2014) who showed that body motion perception needs an integration of multiple visual processes involving Gestalt-like pattern and aggregation of the bottom-up and up-down processes with recognition based on learning [90]. Visual motion perception uses the integrated dynamic motion model to handle diverse moving stimuli involving motion integration and detection to perceive correctly in decision [91] and can be modelled like [92] or linked between the patterns and perception duration into three groups according to their direction cues, namely, cardinal, diagonal, and toward diagonal [93].

The link between imagery and perception was investigated by putting the observers in the dark, which rotates in the left or right. The velocity of chair rotation should be high, and the direction of imagined rotation is different from physical rotation [94], this motivated suppression of surrounding spatial information which seems losing information [95]. Matsumoto et al. (2015) analyzed schizophrenia patients who have impairments in cognition, perception, and visual attention; they also analyzed the biological motion perception in 17 patients and 18 healthy controls [96]. Ahveninen et al. (2016) investigated the combination of spatial and non-spatial information in the auditory cortex (AC) of two parallel streams, namely, “what” and “where” that are modulated for visual cortex subsystems, as well as their integration regarding object perception. This approach uses animated video clips of two audiovisual objects, namely, black and gray cats, and records the magneto- and electroencephalography data. The events in sound are initially linked to object perception in posterior AC, with modulation representations in anterior AC [97].

4.4. Summary

Action perception methods are dominant in computational models for recognition of actions and different techniques used to computationalize this concept. Motion patterns methods (e.g., [10,17]) involved recognition of complex motions using global and local features, whereas more partially focus on locomotion of human subject (forms or motions no discrimination). Kinetic–geometric models

(e.g., [17,18]) focus more on the shape of the actor and make connection between MLD and actor shape description. These models are still popularly used due to their application as object recognition task and even recently more developed by the applications of deep learning method (such as [98,99]). Kinetic–geometric models also used for modification of the methods involving “what” and “where” pathways and in general models with two separated pathways and improved form pathway (answer to “what”). Spatiotemporal filters often used to replace the MLD ([15,16,38,39,100–106]) and expanded into 3D structural methods ([1,51]), even new deep learning approaches involving 3D recognition of human action (i.e., [99]). Table 1 summarizes the pros and cons of action recognition approaches.

Table 1. The perception approaches presented with their contribution in the field.

Approaches in Perception	Topic of Each Approach	Connection to Other Researches
E. J. Marey and E. Muybridge (1850s)	moving photographs presenting locomotion	
Rubin (1927)	visual perception of real movement	
Duncker (1929)	visual perception of real movement	
Johansson et al. (1973)	motion patterns for humans& animals as biological motion (MLD)	
Turaga et al. (2008)	locomotion analysis	
Johansson (1975)	perception of human motion in neuroscience analysis	
Kozlowski & Cutting (1977)	with females and upper body MLD, gender recognition	
Marr et al. (1978)	computational process in the human visual system, 3D shapes	perception
Perrett et al. (1985)	the temporal cortex of macaque monkey analysis found two cells in the brain sensitive for rotation and view of the body movements	
Perrett et al. (1989)	view centered, view independent responses among the brain cells	
Goddard (1989)	spatial and temporal feature incorporation through diffuses MLD data	perception-computer
Goddard (1992)	synergistic manner of the process of “what” and “where” in visual system	neuroscience
Goodale & Milner (1992)	projection perceptual information from striate and inferotemporal cortex	neuroscience-object
Cédras and Shah (1995)	motion-based recognition into motion models	identification
Perkins (1995)	animated real-time, texture of motion, avoiding computational	modelling
Tsai et al. (1993)	detection of cyclic motion, applying Fourier transform	modelling
Fielding & Ruck (1995)	Hidden Markov Model (HMM) technique for classification	highly related to
Gallese et al. (1996)	analysis the electrical activity in macaque monkey’s brain	computational modelling
Aggarwal et al. (1999)	human motion analysis review and computer vision approaches	neuroscience
Aggarwal & Cai (1997)	interpreting human motion, tracking, recognizing human activities per frame	computer vision
Decety & Grèzes (1999)	Process of action and its perception, functional segregation MLD	perception
Rangarajan et al. (1992)	matching the biological motion trajectories (object recognition)	
McLeod (1996)	motion blind patient, the homologue of V5/MT concerning the moving stimuli	computer vision
Wiley & Hahn (1997)	virtual reality approach regarding the computer-generated characters	
Neri et al. (1998)	visual system ability to integrate the motion information of walkers	
Hill & Pollick (2000)	temporal differences in MLD, recognition of the exaggerated motions	
Giese & Poggio (2000)	linear combination of prototypical views,3D stationary object recognition	computer vision
Moeslund & Granum (2001)	a comprehensive survey on the motion capture	computer vision
Grèzes et al. (2001)	neural network specifications and its verifications through fMRI	computer
Pollick et al. (2001)	visual perception effects used point-light display(MLD)	vision-neuroscience
Song et al. (2001)	Computer-human interface using joint probability density function (PDF)	computer vision
Servos et al. (2002)	relationship between biological motion and control unpredicted stimuli	computer vision
Grossman & Blake (2002)	neural mechanisms, anatomical, and functionality into two pathways	computer
Jastorff et al. (2002)	investigating of recognition process in the neural mechanism	neuroscience
Giese & Lappe (2002)	spatio-temporal generalization of the biological movement perception	neuroscience
Beintema & Lappe (2002)	analysis of the perception of form pattern of human action by MLD	computer vision
Puce & Perrett (2003)	analysis of single cells, neuroimaging data and records of field potential	
Kilner et al. (2003)	analysis of the action in motor programs	neuroscience
Cohen & Singh (2006)	perceiving the complex shape orientation and local geometric attributes	neuroscience
Lange et al. (2006)	moving human figure using MLD	neuroscience
Troje & Westhoff (2006)	data retrieval of direction from scrambled MLD in humans and animals	
Blake & Shiffrar (2007)	review in perception	
Pyles et al. (2007)	comparative research on human MLD	
Gölcü & Gilbert (2009)	perception of object recognition analyzing the human features	neuroscience
Daems & Verfaillie (2010)	analysis of body postures is different viewpoints and human identification	
Grossman (2010)	analyzing STSp region and its functionality underlying the BLOD response	modelling
Strasburger et al. (2011)	Peripheral vision and pattern recognition for theory of form perception	neuroscience
Giese (2014)	complex pattern recognition mechanism and motion perception	modelling
Tlapale (2015)	integrated dynamic motion model (IDM) to handle diverse moving stimuli	
Jung & Gu (2015)	perception and modeling in the visual motion	modelling
Meso & Masson (2015)	characterized the patterns and perception duration,	
Nigmatullina (2015)	link between the imagery and perception	
Tadin (2015)	spatially suppress the surrounding by perception information	
Matsumoto et al. (2015)	analyzed the biological motion perception in Schizophrenia patients	
hveninen et al. (2016)	combination of spatial and non-spatial information in auditory cortex (AC)	neuroscience

5. Knowledge Based Modeling Approaches

Modeling of biological movements into systematical and mathematical models follows the neuro-physiological, physiological, and neuro-science evidence. This modeling is increasingly developed and considered one branch of this research field; many computer vision approaches

also underlie this model. Sometimes engineering approaches have mixed into such modelling methods such as using HMM and features based bottom-up approaches in time sequence images [11,107,108], however, some of them can be interpreted in the perspective of biologically inspired approaches (e.g., [75,98,104,109–113]). We divide the methodology in these approaches into several directions based on their popularity, which followed neurophysiological evidence discussed in the previous sections. Figures 3 and 4 represent computational model proposed based on biological evidence.

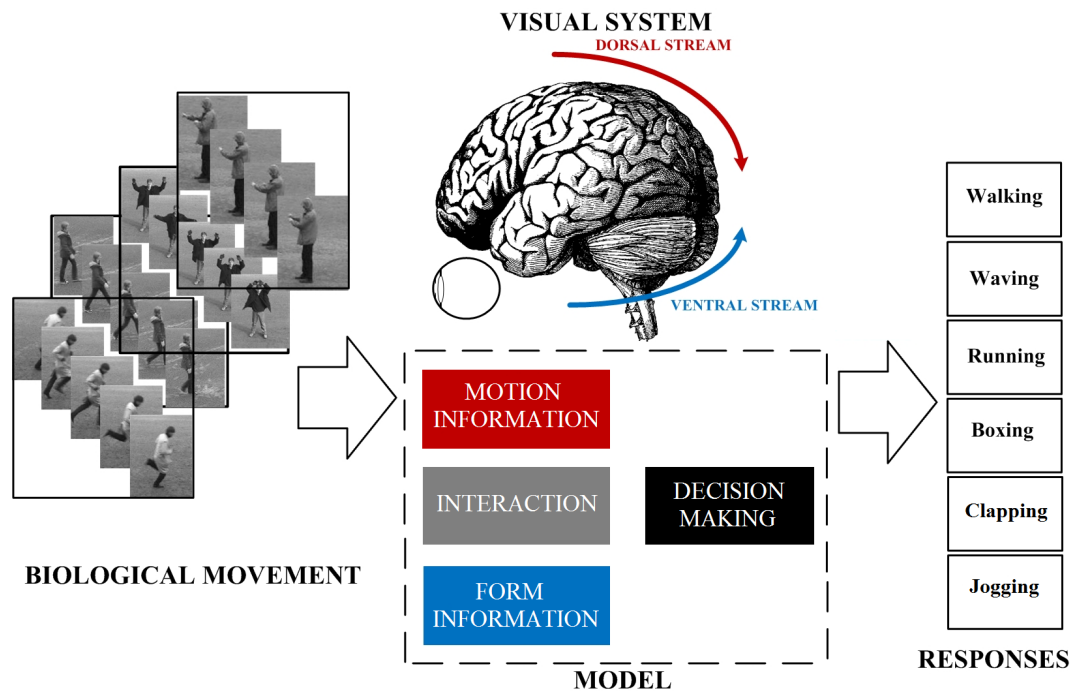


Figure 4. An overall trend of biologically inspired model for human action recognition overview diagram. The two parallel processing streams are considered for the form and motion information.

5.1. Gabor Filter in Form Pathway

Furthermore, Webb et al. (2008) presented the mechanism in intermediate levels of visual processing and investigation to detect circular and radial forms. This mechanism analyzes the detection of the global structure in spiral form using the array consisting of 100 Gabor that is randomly positioned within the window. The Gabor filter randomly rotates, and the structure can be detected when the mask and test have the same spiral pitch (Figures 3–5). The Gabor filter is extensively used in the form pathway, and the approach is significant for elucidating the mechanism of visual processing streams [114]. The Gabor filter (assumed as zero-centered) is the product of a sinusoid and a Gaussian and can be presented by following formula:

$$GW(x, y; \lambda, \theta, \phi, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \phi\right), \quad (1)$$

where $x' = x \cos \theta + y \sin \theta$, $y' = -x \sin \theta + y \cos \theta$. Let wavelength is the number of cycles/pixel is λ . Orientation is the angle of the normal to the sinusoid is θ and phase is the offset of the sinusoid is ϕ . Aspect Ratio is showed by an ellipticity is produced with $\gamma < 1$. Some studies investigate bio-inspired models of human action recognition that focus more on the influence of spiking neural networks in the visual cortex [15,38,101,102,115]. Shu et al. (2014) used a 3D Gabor filter was tailored for V1 cells based on a hierarchical architecture and considered two visual cortical areas, namely, middle temporal area (MT) and primary visual cortex (V1) for motion processing [115]. This model is modified to be a supervised Gabor-based object recognition approach called ABM [15] in the ventral

processing stream) [38,102] (Figure 5a,c). A fuzzy-based optical flow proposed for dorsal streams was used to improve the model [100,101]. Furthermore, an approach to involve slow features was presented for ventral processing stream [16,103]. The active basis model [15] that applies Gabor wavelets (for the elements dictionary) consists of a deformable biological template. A Shared Sketch Algorithm (SSA) is followed through AdaBoost.

$$I = \sum_{i=1}^n c_i \beta_i + \epsilon \quad (2)$$

where $\beta = (\beta_i, i = 1, \dots, n)$ is set of Gabor Wavelet elements and components of sin and cosine, $c_i = \langle I, \beta_i \rangle$ and ϵ is unsolved image coefficient [15,103]. A Gabor wavelets dictionary, comprising n directions and m scales is in the form of, $GW_j(\theta, \omega), j = 1, \dots, m \times n$. Where, $\theta \in \{\frac{k\pi}{n}, k = 0, \dots, n - 1\}$ and $\omega = \{\frac{\sqrt{2}}{i}, i = 1, \dots, m\}$. To find the object's shape, GW features used with small size, location, and posture variance, which concludes the overall shape structure and stays throughout the recognition process. Response (convolution) to every element offers form information with θ and ω .

$$B = \langle GW, I \rangle = \sum \sum GW(x_0 - x, y_0 - y : \omega_0, \theta_0) I(x, y). \quad (3)$$

GW_j is a $[x_g, y_g]$, I is the $[x_i, y_i]$ matrix, and the response of I to GW is $[x_i + x_g, y_i + y_g]$ (with zero-padding). $\{I^m, m = 1, \dots, M\}$ denotes the obtained training set, and B_i is chosen by the joint sketch algorithm. The objective is to identify B_i so that its edge segments obtained from I_m become maximal [15]. It is then necessary to compute $[I^m \cdot \beta] = \psi | \langle I^m \cdot \beta \rangle |^2$ for different i where $\beta \in Dictionary$ and ψ represents sigmoid, whitening and thresholding transformations (β maximizes $[I^m \cdot \beta]$). Let the template; for every training image be $\beta = (\beta_i, i = 1, \dots, n)$, and I^m scores as below:

$$M(I^m, \theta) = \sum_{i=1}^n \delta_i | I^m \cdot \beta | - \log \Phi(\lambda \delta_i). \quad (4)$$

M shows the match scoring function, Φ is a nonlinear function, and δ_i from $\sum_{n=1}^M [I^m, \beta]$ performs steps selection. Weight vectors are calculated by the maximum likelihood technique (by $\Delta = (\delta_i, i = 1, \dots, n)$).

$$Max(x, y) = \max_{(x,y) \in D} M(I_m, \beta). \quad (5)$$

The maximum matching score is calculated by $Max(x, y)$, where D represents the lattice of I , and used for recognition of object recognition in form pathway [15,16,102].

5.2. Deep Learning

In addition, a fully automatic system for human action recognition is presented using convolutional neural networks (CNNs) in the uncontrolled environment. CNN is a deep learning approach, which is bio-inspired and develops a 3D CNN for the task. This approach extracts the features from spatial and temporal dimensions via a convolutional network, captures the motion information encoded from adjacent frames, and generates and combines multiple channel information. The presented approach has successfully implied a bio-inspired method through CNN and motion information combination for actual environments [99] (Another similar approaches are [116–125], long-short term memory (LSTM) [98,104,108,126]).

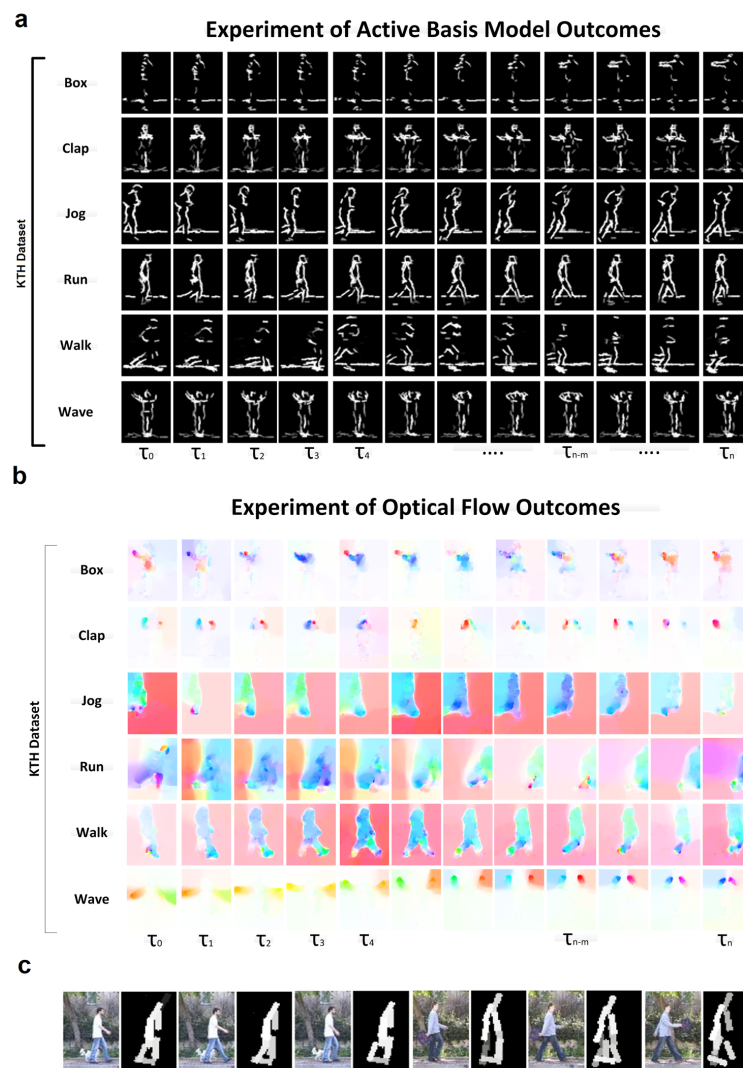


Figure 5. The results of simulation for a biological movement paradigm using ABM-based IncSFA modelled for form pathway (a,c) and motion pathway (b) by applying optical flow [127], are shown [16]. In (a,c) every row shows the response of ABM as well as the slowness features generated for actions. In (b) rows represent actions and a representation of motion information in false color form.

5.3. Sparse Representation

Due to the nature of sparsity, it has been extensively used for modelling ventral stream [128–131]. Lehky et al. (2011) investigated the characteristic of sparseness selection in the anterior inferotemporal cortex on a large dataset which involved the information on 674 monkey inferotemporal cells for 806 object photographs and the two-way analysis of the responses of the entire neurons in single image (population sparseness) and column-wise (response of single neurons to all images). This approach also represents inconsistent structural-based object recognition tasks, and the objects are decomposed into small standard features [128]. Complex visual understanding Lobula giant movement detectors (LGMD) and directional selective neurons, in visual pathways of locusts involved a model that tunes these two networks for collision tasks, compares them separately, and analyzes them co-operationally [129]. A sparse representation is shown in a set of overcompleted basis (dictionary) on different actions includes vector quantization (VQ) and clustering [132]. A non-negative sparse coding is used to learn the basic patterns of motion [133]. Nayak & Roy-Chowdhury (2014) presented an approach using spatiotemporal features and their unsupervised relationship to dictionary learning in the model of activity recognition. This approach provides an unsupervised sparse decomposition

framework for the relationship between the spatiotemporal features and the local information from descriptors, which create classifiers through multiple kernel learning [134,135].

5.4. Dynamic Representation Of Action

The dynamic representation of action recognition was analyzed by different approaches to track the dynamics representation of movements rather than shape of the approaches [109,112,113,136–139]. Some of these approaches involved the human silhouette such as [136], which used a pose descriptor called histogram of oriented-rectangles to represent the human action recognition in the video streams. Such approaches require more consistent connection between the frames of the stream, such as bag-of-word (BOW) representation and time dependent techniques [109,110,137,138,140]. Sometimes BOW mixed up with the basis of vector quantization (VQ)[140], maxima of the sparse interest point operators [130], or Sparse coding [131] to improve the performance of model.

Optical flow is widely used in the computational model to provide motion information (i.e., the layer-wise optical flow) [15,16,38,100–103,141,142]. M_1 and M_2 are visible masks for two frames $I_1(t)$ and $I_2(t - 1)$, and the field of flow from I_1 to I_2 and I_2 to I_1 are represent by (u_1, v_1) , (u_2, v_2) . The following terms will be considered for layer-wise optical flow estimation. The objective function consists of summing three parts; visible layer masks then match to these two images using a Gaussian filter and are called data term matching $E_\gamma^{(i)}$, symmetric $E_\delta^{(i)}$, and smoothness $E_\mu^{(i)}$.

$$E(u_1, v_1, u_2, v_2) = \sum_{i=1}^2 E_\gamma^{(i)} + \rho E_\delta^{(i)} + \xi E_\mu^{(i)}. \quad (6)$$

In the next section, a summary of knowledge based modeling approaches and their approaches to combine the motion and form information is briefly summarized (Figure 5b).

The approach was considered by Marc Jeanerod as the basic method of action (semantics and pragmatics) and movement. The ordinary representational resources of pragmatics and semantic types of actions following the evidence of simulation and language understanding were investigated. Three theoretical frameworks were mentioned by Prinz (2014):

1. Semantics is based on pragmatics;
2. Pragmatics is anchored on semantics; and
3. Pragmatics is a part and parcel of semantics (taken from [143]).

This approach analyzes adaptive local space time features that are captured in the local events located in the video [144]. A computational model follows the neural plausibility assumptions for the interaction of the form and motion signals in biological motion perception from figural form cues; the receptive fields in the images of a static human body were also analyzed [145].

5.5. Interaction Between Pathways

Having two pathways involves the system with two types of data and for making decisions, the system usually combined these information and played an important role [146]. There are several methods proposed for interaction between pathways based on biological evidence (e.g., quantitative models and neurophysiologically plausible tools for model establishment [26]) (Figures 3c and 5–7). The dominant form and motion (optic flow) feature extraction in the mid-level of the moving subject is conducted by using principal component analysis (PCA) from spatial localization are considerably effective to recognize biological movement [24] (has similarity with [147]). Another skeletal representation of human body as the articulated interconnections incorporated a dual square-root function (DSRF) descriptor, where it decomposed the skeleton into five parts [111] and a relevant hierarchy that processes them and provides position-invariant feature detection [106]. An interactive approach was used that estimates motion using optical flow through a dynamic Bayesian network and involves interaction with motion information from two-filter inference in online and offline parameter optimization [142].

A bio-inspired feed-forward of spiking network model was performed for the influences of the motion system (V1 and MT) on human action recognition. Two characteristics of neural code, namely, neuron synchronization and their firing rate, were considered. Spiking networks can be a potential alternative in actual visual applications [148]. Guthier et al. (2014) studied the interaction and combination of the pathways in the visual system. The investigation focused on the recognition of complex biological articulated movements. Figure 7 represents an general overview on computational model, which potentially a representation of advancements in the pathways. They introduced a model that utilizes gradient and optical flow. The patterns are used by an unsupervised learning algorithm, translation-invariant nonnegative sparse coding called VNMF, and shaped prototypical optical flow patterns. In the learning processes, a lateral reserve term that eliminates competing pattern activations provide small sparse activations [149]. An interaction of motion and spatial information was presented using Motion Binary Pattern (MBP) and Volume Local Binary Pattern (VLBP) with Optical Flow for recognizing human actions and space-time volume binary patterns [150]. A fuzzy multiplication of form and motion information is proposed by combining active basis model with optical flow [38,100,101].

Moreover, a combination of slowly varying features and fast varying features proposed by Yousefi et al. (2016) [103]. Haghighi et al. (2016) studied human-like movements processed in the human brain and motor control. The study involved the concept of artificial intelligence and robotics, as well as learning the latent simple motions for imitation in more complex movements. It proposed MOSAIC structure in motor control modeling [151]. An interaction of form and motion pathways using representations of both paths for each selected key pose frame $I_{key} \in K$, a form representation of spatial derivatives combination and concatenating orientation selective maps with following formula:

$$I_{form} = \frac{\log(1 + 5|I^{con}|)}{\max(\log(1 + 5|I^{con}|))} \quad (7)$$

Replacing I^{con} with optical flow I^{flo} (using vertical and horizontal components instead of spatial resolution) and concatenating them and making similar formula for form information as follow:

$$I_{motion} = \frac{\log(1 + 5|I^{flo}|)}{\max(\log(1 + 5|I^{flo}|))} \quad (8)$$

Then, form and motion information are combined as an input to train of the deep convolutional neural network (DCNN) [119,152].

Ward et al. (2010) investigated the reference frames applied in terms of visual information by using fMRI. The analysis considers the receptive field scene processing areas, such as transverse occipital sulcus (TOS), retrosplenial complex (RSC), and parahippocampal place area (PPA). PPA and TOS show the position response curves on the fixation points to the screen (or the pattern), whereas RSC area does not [153].

Bio-inspired features in action recognition are presented by involving the motion in the models of cortical areas V1 and MT (shape and characteristics of their receptive field). A model with different surrounding geometries for MT cell receptive field is presented, which leads to bio-inspired features regarding the average activity of MT cells and how these features are used as a standard in the classification of activity recognition [154].

5.6. Summary

Computational mechanism for action recognition follows a dual pathways model, which involves form and motion information in two independent pathways. Form pathway corresponds to shape information and usually involves object recognition task having Gabor like (spatiotemporal) features motivated by simple and complex cells to grasp the object shape [15,38,100,102,155] (showed in Figure 6). Current developments in deep learning have considerable impact on improvement of this

pathway and researchers still trying to modified this pathway modifying deep learning configuration in a bid of increasing the accuracy or robustness of current approaches. For this pathway, deep learning showed a significant performance using different configurations [116–121,156]. On the other pathway, motion information extracted by optical flow, which was considered as fast varying features [103]. Interaction between dual pathways considered to be a challenge between recognition and space-time dependent patterns, which usually followed the neuro-physiological, physiological evidence rather than heuristic techniques. Table 2 summarizes knowledge based approaches.

Table 2. The Knowledge based modeling approaches presented with their contribution in the field.

Psychological and Neuroscience Approaches	Topic of the Approach	Connection to Other Researches
Jellema et al. (2000)	analysis of the cellular population located in the temporal lobe of the macaque monkey	Psychology
Billard et al. (2000, June)	action imitation considered the actions high-level abstractions	
Vaina et al. (2001)	investigation regarding the neural network, fMRI in MLD	
Goodale & Westwood (2004)	evaluating the labour division at visual pathways	
Banquet et al. (2005)	associative learning for object location level, in CA3-CA1 region	
Milner & Goodale (2006)	involvement of dorsal stream in movement to target following ventral stream	
Cook et al. (2009)	ASCs for comparing detection of non-biological and biological motion	
Wyk et al. (2012)	action representation at STS	Psychology
Schenk (2012a)	DF patient analyzes the ability to get the object	
Schenk (2012b)	Using fMRI the functionality of DF patient	
Fleischer et al. (2013)	visual recognition from motion	
Hesse & Schenk (2014)	visuomotor performance of a D.F. patient tested for letter-posting task	
Theusner et al. (2014)	motion energy based on the luminance of objects motion detectors	
Whitwell et al. (2014)	test by different width of the objects and DF, distinguish the shape perceptually	perception
Whitwell et al. (2015)	ability grip scaling is may rely on online visual or haptic feedback (for DF patient)	
Krigolson et al. (2015)	review of the behavior using EEG	Psychology
Cavina-Pratesi et al. (2015)	brain circulation using fMRI regarding the word recognition ability	
Ganos et al. (2015)	voluntary movement considering GTS area	
Yamamoto & Miura (2016)	visual object motions on time perception	
Schindler & Bartels (2016)	on 3 dimensional visual cue involving the motion parallax analyzing	
Venezia et al. (2016)	the sensorimotor integration of visual speech through the perception	perception
Harvey & Dumoulin (2016)	visual motion effects on neural receptive field and fMRI response	

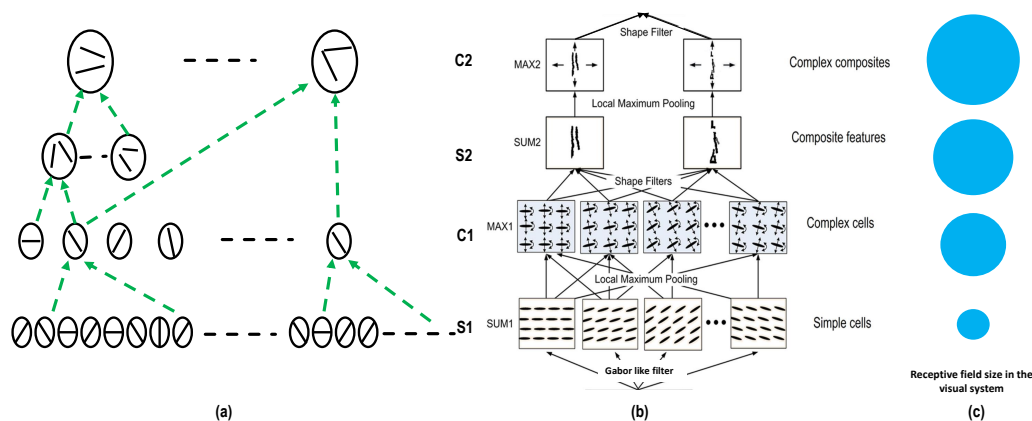


Figure 6. (a) The scheme representation of a stream of visual processing located in human cerebral cortex and the ventral stream information of early visual areas (V1p) and its projection in the occipito-temporal cortex and dorsal stream (blue) projection of the information to the posterior parietal cortex are shown. The indicated routes represented by the arrows and involvement of complex interconnections [155]; (b) also follows the scheme representation of human (or any objects) in ventral stream from simple cells to complex composites; (c) a presentation of receptive field from early vision till recognition in both pathways.

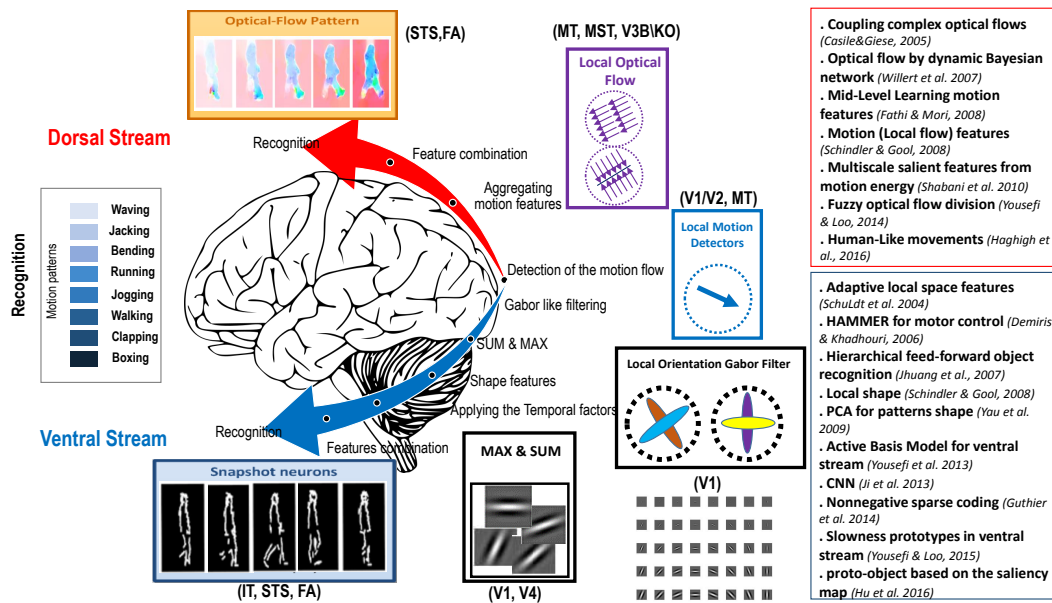


Figure 7. A brief visual representation of some important approaches in modifying computational model for human action recognition.

6. Psychological and Neuroscience Point Of View

6.1. Biological Evidence Using Fmri

The cellular population located in the temporal lobe of macaque monkeys' inner superior temporal sulcus (partially called STPa or STSa) was analyzed. The responses of these cells were associated with the agent's action performance as it reached the targeted position [157]. Some other types of actions (oscillatory movements, repetitive form of the arms and legs, and exact movements of reaching and grasping) investigated in [158], which was expanded by fMRI and point light research involving rigid and non-rigid motion responses of biological motion with gender concerns [90,159–161] (Reference [90] presented body posture for spatiotemporal receptive fields, Reference [160] studied hand actions, and Reference [161] presented an engineering approach but inspired by biological evidence). An investigation using fMRI for 37 children showed both groups of gender bilaterally showed the activation at the posterior STS; in response to incongruence, children showed response changes in the STS regions, also an incongruency effect was observed in the older children and adolescents in the experiment [162] (also similar to [163]). Goodale and Westwood (2004) presented another approach that evaluates the labor division at visual pathways and completed their hypothesis on the primate cerebral cortex between ventral streams dedicated to visual perception and dorsal stream for visual control in action. The study analyzed the psychological evidence and the response to visual motor control; in particular, the neurobiological challenge in mapping these behavioral findings onto the brain was analyzed and compared with known information about ventral and dorsal streams (in primate neurophysiology and human neurophysiology) [155]. An associative learning type analysis was presented for object location level, spatiotemporal level in the CA3–CA1 region, and movement-related information in the entorhinal cortex. This letter also analyzed the behavioral implementation and multi-modal integration, which suggested the functional interpretation in hippocampo-cortical systems [164]. Yamamoto and Miura (2016) analyzed visual object motions on time perception. They investigated the line segments in front or behind the occluders in different speeds and followed the association of time perception with global motion processing [165]. A study on 3D visual cue involving motion parallax analyzed the link between the visual motion and scene processing by using fMRI. Parallax-selective responses were found in parietal regions IPS4 and IPS3, and in the region of occipital place area. Some regions such as the RSC and PPA do not respond to the

parallax [166]. Venezia et al. (2016) analyzed the sensorimotor integration of visual speech through perception. The study used fMRI on healthy individuals to identify new visiomotor circuit speech production [167]. A research on how visual motion affects neural receptive fields and fMRI response amplitudes was carried out to examine visual motion neural position preferences in the hierarchy of visual field maps using high-field fMRI and population receptive field. The results showed that visual motion induces the transformation of visuo-spatial representations through the visual hierarchy [168].

6.2. Biological Model And Imitation

Another approach on action imitation explains the natural action through visual analysis of actions and motor representation of the nervous system. The evidence of the existing system is mentioned by mirror system for mapping in primates and humans [169]. The human imitation of machines has been investigated for the purpose of flexibility, usefulness, and development of user-friendly machines. The approach concentrates on understanding how robots determine what to imitate, as well as the process of mapping perception onto the action it is imitating [170]. A mathematical approach that tackle parts of the imitation problem and the motor side of the imitation were investigated. The results argued that the perceptual system for movement identification and the spatial information correspond to these actions [171]. The cognitive development agent in the imitation and its architecture in the recognition of action was presented and implemented in the robots. The understanding and generation of actions, as well as the ability to learn new composite actions during the mentioned architecture, were also investigated [172] (similar to [173] another bio-inspired robotic imitation/recognition approach).

6.3. Visual System Impairment and Pathways

A study on autism spectrum conditions (ASCs) compared the detection of non-biological and biological motion in human adults through psychological evidence obtained from participants who watched biological (hand movements) and non-biological (falling tennis balls) stimuli. The ASC group did not show proper responses to perturbation from biological motions based on velocity profile [174]. The dorsal stream was suggested to be involved in movement to the target following ventral stream visual representation processing delay [175]. The visuomotor performance of a DF patient was tested through a letter-posting task. The absence of environmental cues was observed in the DF patient, causing them to be unaffected by delay (aforementioned). The findings suggest that ventral stream damage does not consistently influence delayed movements but affects the visual feedback and environmental landmarks [176]. Another investigation on DF patients analyzed their ability to acquire an object by distinguishing its geometry. Using fMRI, the functionality of a DF patient uses the intact visuomotor system housed within the posterior parietal lobe in the dorsal stream. Moreover, Schenk (2012a, 2012b) described the non-functioning of visuomotor networks in the dorsal stream, which was caused by a haptic feedback of the targeted object's edges [177,178]. A test was conducted using different object widths, and the DF patients could grasp them within the healthy range (unlike the hypothesis that they should not); moreover, haptic feedback did not improve the ability of the DF patient to distinguish the shape perceptually [179]. Another research mentioned that ability grip scaling may rely on online visual or haptic feedback. The grip scaling of the DF patient did not activate while her vision was suppressed in a grasp movement, showing that the haptic feedback after perception impairs the DF patient's performance. The research showed that DF patient's spared grasping task relies on dorsal stream functioning at the normal mode [180]. Krigolson et al. (2015) presented a review of the behavior on three areas, namely, feedback processing, feed-forward control, and target perturbation; electroencephalography (EEG) was utilized to determine the temporal nature application in the goal-directed action. The cognitive potential and neural processing timing to motor control were further analyzed [181]. A research used fMRI to investigate brain circulation during word recognition in the left fusiform and left inferior frontal areas of the gyrus, as well as the left middle temporal cortex of the DF patient. The left fusiform activations called visual word form area appeared from the FFA and hypothesized that this area lies outside the LO [182]. A research on voluntary

movement in the Gilles de la Tourette syndrome comprised 25 patients. The results suggested that the brain learns voluntary control by perceptually discriminating signals from noise [183].

6.4. Summary

Neurophysiological evidence shows the connection between different parts of the brain during visual recognition of motion. Despite there being many studies in this area, this field is still considered to be an action research direction in a bid to develop computational neural mechanism [184–186]. The impairments in visual system reported in the evidence proved the concept of dual pathways (i.e., form and motion). Moreover, evidences gathered by fMRI and EEG continues to modify current developments in the computational model. Table 3 and Figure 7 summarize the psychological and neuroscience approaches along with targeted areas of the brain with their computational tool in the model.

Table 3. The psychological and neuroscience approaches presented with their contribution in the field.

Knowledge Based Modeling Approaches	Topic of the Approach	Connection to Other Researches
Yamato (1992,June)	HMM and feature based bottom up approaches in time sequence images	modelling
Giese & Poggio (1999)	Linear combination of motion sequence prototypical views, 3D object recognition	
Gavrila (1999)	survey article in visual analysis regarding the human movement	
Wachter & Nagel (1997, June)	the quantitative description of the geometry of the human object	
Gises & Poggio (2003)	dual processing pathways in the visual system	
Schuldt et al. (2004)	adaptive local space-time features	Computer vision
Casile & Giese (2005)	multilevel generalization using simple mid-Level optic flow features	perception
Arbib (2005)	analysis of neural and its functionality grounding for the Language skills	perception, neuroscience
Valstar & Pantic (2006)	comparison of Logical and biological inspired methods for facial expression	computer vision
Demiris & Khadhour (2006)	computational architecture and HAMMER for motor control systems	perception
Lange & Lappe (2006)	Neural plausibility assumptions for interaction of the form and motion signals	
Willert et al. (2007)	estimating the motion using optical flow by dynamic Bayesian network	
Jhuang et al. (2007)	hierarchical feed forward architecture on the object recognition	
Minler and Goodale (2007)	analysis of two cortical systems regarding the vision in action	perception
Fathi & Mori (2008)	mid-Level Learning for the motion features	modelling
Schindler & Gool (2008, June)	recognition of simple actions instantaneously by short sequences (snippets) 1-10 frames	computer vision
Schindler & Gool (2008)	recognition of form (Local shape) and motion (Local flow) features	computer vision
Webb et al. (2008)	intermediate Levels of visual processing, detection circular, and radial form	
Willert & Eggert (2009)	estimation of motion to analyze the small number of temporal consecutive frames	
Yau et al. (2009)	interaction of vision and touch, PCA for patterns shape features identification	
Escobar et al. (2009)	bio-inspired feed-forward of spiking network model	neuroscience
Ikizler & Duygulu (2009)	analyzing the dynamic representation of action recognition using HOR	
Wang & Mori (2009)	visual features as visual word and semi-Latent topic models	modelling
Ryoo & Aggarwal (2009)	Spatiotemporal relation for recognition of human activity	
Dean et al. (2009, December)	Learning sparse spatiotemporal codes from the basis vectors	
Shabani et al. (2010)	multiscale salient features from motion energy	modelling
Poppe (2010)	Visual based human action recognition	computer vision
Sun et al. (2010)	median filtering of the intermediate flow fields	
Ward et al. (2010)	references frames applied for visual information using fMRI	
WeinLand et al. (2011)	review paper for human action/activity recognition	
Lehky et al. (2011)	characteristic of selection of sparseness in anterior inferotemporal cortex	neuroscience
Willert & Eggert (2011)	representation of visual motion processing	
Mathe & Sminchisescu (2012)	BOW in maxima of sparse interest operators	
Escobar & Kornprobst (2012)	analysis motion in the models of cortical areas V1 and MT	neuroscience-perception
Cadiou & Olshausen (2012)	intermediate-level visual presentation	
Guha & Ward (2012)	human action in the sparse representation in overcompleted basis (dictionary) set	
Guthier et al. (2012)	non-negative sparse coding on biological motion	
Yousefi et al. (2013)	Introducing Active Basis Model for ventral stream	Computer vision
Ji et al. (2013)	fully automatic system for human action recognition by CNN	modelling
Pitzalis et al. (2013)	motion analysis approach considers the movements in all directions	perception
Yue & Rind (2013)	detection of collisions, analysis of two types neurons: LGMD and DSNs	neuroscience
Cai et al. (2014)	spatiotemporal feature in the bio-inspired model, BIM-STIP	
Guthier et al. (2014)	survey, modelling using nonnegative sparse coding, VNMF	
Shu et al. (2014)	bio-inspired modeling human action recognition, spiking neural network	
Nayak & Roy-Chowdhury (2014)	spatiotemporal features, unsupervised way into a dictionary Learning	
Yousefi & Loo(2014)	fuzzy optical flow division in Dorsal stream	Computer vision
Esmaili et al. (2014)	robust recognition of face using C2 features in HMAX	
Yousefi & Loo(2014)	Interaction between dorsal and ventral streams	Computer vision
Prinz (2014)	analysis of action semantics and pragmatics	perception
Yousefi & Loo(2015)	Slowness principal into modeling	Computer vision
Moayedi (2015)	basic shape extraction of action group sparse coding employed BOW	
Yousefi & Loo(2015)	Hybrid Max-Product Neuro-Fuzzy Classifier and Quantum-Behaved PSO in the model	Computer vision
Tian (2015)	BOW method, VQ ,CLLC, GSRC in the human action recognition	
Yousefi & Loo(2015)	Slowness prototypes in the ventral stream	Computer vision
Hu et al. (2016)	proto-object based on the saliency map	computer vision
Haghigh et al. (2016)	human-Like movements	

7. Future Directions

Several open problems should be solved to allow us to develop the methodology further. Here, we discuss some future directions in biologically inspired action recognition that might be interesting to explore.

Neurophysiological evidences: One direction involves further investigation of neuroscience, and neurophysiological evidences for a better perception of the biology of the brain. Evidently, the presented approaches carefully follow the existing evidence in the field, and further framework requires explicit details in biological studies of visual system, which can initiate further developments in the mechanism. (i.e., [187–189])

Development in two pathways methodology: The current frameworks include different methods to cover the requirements for the model with two pathways. The computational load of these combinations should also be considered. This step would allow adapting or extending into more complex analyses, which provide advancements in the current mechanism, which involves structural configurations or methodological developments (i.e., [190]).

Deep learning and the biologically inspired mechanism: Finally, in terms of machine learning, another possibility would be to create another machine learning framework (with respect to biological evidences) and modify the system from episodic recognition or the frame recognition with overall understanding of the movements. In this area, with the recent developments in deep learning approaches, this concept is implemented and would be a good methodology in term of involving shape features (from form pathway) and motion information (from motion pathway) (some examples are deep learning applications, which can be justified with biological connection [98,99,104,108,116–126]). One particular trend is on applying the framework to learn more complexity in biological movements depending on deep learning based machine vision applications.

8. Conclusions

The presence of neuroscience and neurophysiological evidence as motivating facts for modifying models changes the research focus about human action recognition from a computer vision method to computational neural mechanism. We have presented a relatively complete survey of state-of-the-art methods for biologically-inspired action recognition. The reviewed techniques merged several emerging fields focusing on different perspectives of biological movements. We gathered several aspects of such perspective involving action perception, computational and knowledge based modeling, psychological, and neuroscience approaches in the field along with providing future research directions.

Funding: This research was partially funded under the High Impact Research (HIR) foundation grant: contract No.UM.C/HIR/MOHE/FCSIT/10, at University of Malaya (UM), Malaysia.

Acknowledgments: The authors tried to consider the state-of-the-art approaches and a large set of scientific papers where, for the sake of accessibility for the reader, have been collected and we have largely restricted ourselves to the relevant papers. There was a selection inevitably and hence some papers have been omitted as this is the nature of any review. We apologize to those authors whom their papers have not been included. Also, we would like thank anonymous reviewers for their insightful and constructive comments.

Conflicts of Interest: Authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	artificial intelligence
MLD	moving light display
PGA	Parliamentarians for Global Action
HMM	hidden Markov model
2D	two dimension
3D	three dimension

STS	superior temporal sulcus
ANN	Artificial Neural Network
KO	Kinetic Occipital area
MT	Medial Temporal
V5	Fifth portion of visual extrastriate areas
LOC	loss of consciousness
EBA	extrastriate body area
BOLD	blood oxygenation level-dependent
fMRI	functional magnetic resonance imaging
STSp	posterior superior temporal sulcus
ITS	inferior temporal sulcus
FFA	fusiform face area
FBA	fusiform body area
AC	auditory cortex
DSRF	Dual Square-Root Function
HAMMER	multiple models for execution and recognition
ANN	Artificial Neural Network
PPA	parahippocampal place area
TOS	transverse occipital sulcus
RSC	retrosplenial complex
V1	Visual Primary Cortex
CNNs	convolutional neural networks
LSTM	long-short term memory
LGMD	Lobula giant movement detectors
STIP	spatio-temporal interesting points
MBP	Motion Binary Pattern
VLBP	Volume Local Binary Pattern
OPE	optical flow field
BOW	bag of visual words
MST	Medial Superior Temporal
V3A	Third area of visual extrastriate areas-accessory
CSv	Cingulate Sulcus Visual Area
IPSmot	Intra-Parietal Sulcus motion
ABM	active basis model
DRAMA	Dynamical Recurrent Associative Memory Architecture
ASCs	autism spectrum conditions
LGNd	lateral geniculate nucleus in the thalamus
V1p	early visual areas
EEG	electroencephalography

References

1. Aggarwal, J.K.; Cai, Q. Human motion analysis: A review. In Proceedings of the Nonrigid and Articulated Motion Workshop, San Juan, PR, USA, 16 June 1997; pp. 90–102.
2. Turaga, P.; Chellappa, R.; Subrahmanian, V.S.; Udrea, O. Machine recognition of human activities: A survey. *IEEE Trans. Circuits Syst. Video Technol.* **2008**, *18*, 1473–1488. [[CrossRef](#)]
3. Rubin, E. Visuell wahrgenommene wirkliche Bewegungen. *Z. Psychol.* **1927**, *103*, 384–392.
4. Duncker, K. Über induzierte bewegung. *Psychol. Forsch.* **1929**, *12*, 180–259. [[CrossRef](#)]
5. Johansson, G. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* **1973**, *14*, 201–211. [[CrossRef](#)]
6. Leek, E.C.; Cristino, F.; Conlan, L.I.; Patterson, C.; Rodriguez, E.; Johnston, S.J. Eye movement patterns during the recognition of three-dimensional objects: Preferential fixation of concave surface curvature minima. *J. Vis.* **2012**, *12*, 7. [[CrossRef](#)]
7. Santofimia, M.J.; Martinez-del Rincon, J.; Nebel, J.C. Episodic reasoning for vision-based human action recognition. *Sci. World J.* **2014**, *2014*, 270171. [[CrossRef](#)]

8. Hogg, T.; Rees, D.; Talhami, H. Three-dimensional pose from two-dimensional images: A novel approach using synergetic networks. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; pp. 1140–1144.
9. Schindler, K.; Van Gool, L. Action snippets: How many frames does human action recognition require? In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.
10. Schindler, K.; Van Gool, L. Combining densely sampled form and motion for human action recognition. In *Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 122–131.
11. Efros, A.A.; Berg, A.C.; Mori, G.; Malik, J. Recognizing action at a distance. In Proceedings of the Ninth IEEE International Conference on Computer Vision, Nice, France, 13–16 October 2003; p. 726.
12. Daugman, J.G. Two-dimensional spectral analysis of cortical receptive field profiles. *Vis. Res.* **1980**, *20*, 847–856. [[CrossRef](#)]
13. Olshausen, B.A.; Field, D.J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **1996**, *381*, 607. [[CrossRef](#)]
14. Riesenhuber, M.; Poggio, T. Neural mechanisms of object recognition. *Curr. Opin. Neurobiol.* **2002**, *12*, 162–168. [[CrossRef](#)]
15. Wu, Y.N.; Si, Z.; Gong, H.; Zhu, S.C. Learning active basis model for object detection and recognition. *Int. J. Comput. Vis.* **2010**, *90*, 198–235. [[CrossRef](#)]
16. Yousefi, B.; Loo, C.K. A dual fast and slow feature interaction in biologically inspired visual recognition of human action. *Appl. Soft Comput.* **2018**, *62*, 57–72. [[CrossRef](#)]
17. Johansson, G. Visual motion perception. *Sci. Am.* **1975**, *232*, 76–89. [[CrossRef](#)] [[PubMed](#)]
18. Kozlowski, L.T.; Cutting, J.E. Recognizing the sex of a walker from a dynamic point-light display. *Percept. Psychophys.* **1977**, *21*, 575–580. [[CrossRef](#)]
19. Perrett, D.; Smith, P.; Mistlin, A.; Chitty, A.; Head, A.; Potter, D.; Broennimann, R.; Milner, A.; Jeeves, M. Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: A preliminary report. *Behav. Brain Res.* **1985**, *16*, 153–170. [[CrossRef](#)]
20. Perrett, D.I.; Harries, M.H.; Bevan, R.; Thomas, S.; Benson, P.; Mistlin, A.J.; Chitty, A.J.; Hietanen, J.K.; Ortega, J. Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* **1989**, *146*, 87–113. [[PubMed](#)]
21. Goddard, N.H. The interpretation of visual motion: Recognizing moving light displays. In Proceedings of the Workshop on Visual Motion, Irvine, CA, USA, 20–22 March 1989; pp. 212–220.
22. Jamshidnezhad, A.; Nordin, M.J. Bee royalty offspring algorithm for improvement of facial expressions classification model. *Int. J. Bio-Inspired Comput.* **2013**, *5*, 175–191. [[CrossRef](#)]
23. Babaeian, A.; Babae, M.; Bayestehtashk, A.; Bandarabadi, M. Nonlinear subspace clustering using curvature constrained distances. *Pattern Recognit. Lett.* **2015**, *68*, 118–125. [[CrossRef](#)]
24. Casile, A.; Giese, M.A. Critical features for the recognition of biological motion. *J. Vis.* **2005**, *5*, 6. [[CrossRef](#)]
25. Arbib, M.A. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behav. Brain Sci.* **2005**, *28*, 105–124. [[CrossRef](#)]
26. Giese, M.A.; Poggio, T. Neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* **2003**, *4*, 179–192. [[CrossRef](#)]
27. Goddard, N.H. *The Perception of Articulated Motion: Recognizing Moving Light Displays*; Technical Report; DTIC: Fort Belvoir, VA, USA, 1992.
28. Giese, M.; Poggio, T. Synthesis and recognition of biological motion patterns based on linear superposition of prototypical motion sequences. In Proceedings of the Multi-View Modeling and Analysis of Visual Scenes, Fort Collins, CO, USA, 26 June 1999; pp. 73–80.
29. Goodale, M.A.; Milner, A.D. Separate visual pathways for perception and action. *Trends Neurosci.* **1992**, *15*, 20–25. [[CrossRef](#)]
30. Cedras, C.; Shah, M. Motion-based recognition a survey. *Image Vis. Comput.* **1995**, *13*, 129–155. [[CrossRef](#)]
31. Perkins, D. *Outsmarting IQ: The Emerging Science of Learnable Intelligence*; Simon and Schuster: New York, NY, USA, 1995.
32. Tsai, P.S.; Shah, M.; Keiter, K.; Kasparis, T. *Cyclic Motion Detection*; Computer Science Technical Report; University of Central Florida: Orlando, FL, USA, 1993.

33. Riesenhuber, M.; Poggio, T. Hierarchical models of object recognition in cortex. *Nat. Neurosci.* **1999**, *2*, 1019–1025. [[CrossRef](#)] [[PubMed](#)]
34. Hubel, D.H.; Wiesel, T.N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* **1968**, *195*, 215–243. [[CrossRef](#)] [[PubMed](#)]
35. Gallese, V.; Fadiga, L.; Fogassi, L.; Rizzolatti, G. Action recognition in the premotor cortex. *Brain* **1996**, *119*, 593–609. [[CrossRef](#)] [[PubMed](#)]
36. Tarr, M.J.; Bülthoff, H.H. Image-based object recognition in man, monkey and machine. *Cognition* **1998**, *67*, 1–20. [[CrossRef](#)]
37. Billard, A.; Mataric, M.J. Learning human arm movements by imitation: Evaluation of a biologically inspired connectionist architecture. *Robot. Auton. Syst.* **2001**, *37*, 145–160. [[CrossRef](#)]
38. Yousefi, B.; Loo, C.K.; Memariani, A. Biological inspired human action recognition. In Proceedings of the 2013 IEEE Workshop on Robotic Intelligence In Informationally Structured Space (RiSS), Singapore, 16–19 April 2013; pp. 58–65.
39. Fielding, K.H.; Ruck, D.W. Recognition of moving light displays using hidden Markov models. *Pattern Recognit.* **1995**, *28*, 1415–1421. [[CrossRef](#)]
40. Hill, H.; Pollick, F.E. Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychol. Sci.* **2000**, *11*, 223–228. [[CrossRef](#)]
41. Weinland, D.; Ronfard, R.; Boyer, E. A survey of vision-based methods for action representation, segmentation and recognition. *Comput. Vis. Image Underst.* **2011**, *115*, 224–241. [[CrossRef](#)]
42. Giese, M.; Lappe, M. Measurement of generalization fields for the recognition of biological motion. *Vis. Res.* **2002**, *42*, 1847–1858. [[CrossRef](#)]
43. Ryoo, M.S.; Aggarwal, J.K. Spatio-temporal relationship match: Video structure comparison for recognition of complex human activities. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; pp. 1593–1600.
44. Cai, B.; Xu, X.; Qing, C. Bio-inspired model with dual visual pathways for human action recognition. In Proceedings of the 2014 9th International Symposium on Communication Systems, Networks & Digital Sign (CSNDSP), Manchester, UK, 23–25 July 2014; pp. 271–276.
45. Rangarajan, K.; Allen, W.; Shah, M. Recognition using motion and shape. In Proceedings of the 11th IAPR International Conference on Pattern Recognition, Hague, The Netherlands, 30 August–3 September 1992; pp. 255–258.
46. Neri, P.; Morrone, M.C.; Burr, D.C. Seeing biological motion. *Nature* **1998**, *395*, 894–896. [[CrossRef](#)] [[PubMed](#)]
47. Gavrilu, D.M. The visual analysis of human movement: A survey. *Comput. Vis. Image Underst.* **1999**, *73*, 82–98. [[CrossRef](#)]
48. Wachter, S.; Nagel, H.H. Tracking of persons in monocular image sequences. In Proceedings of the Nonrigid and Articulated Motion Workshop, San Juan, PR, USA, 16 June 1997; pp. 2–9.
49. Blais, C.; Arguin, M.; Marleau, I. Orientation invariance in visual shape perception. *J. Vis.* **2009**, *9*, 14. [[CrossRef](#)]
50. Wang, Y.; Zhang, K. Decomposing the spatiotemporal signature in dynamic 3D object recognition. *J. Vis.* **2010**, *10*, 23. [[CrossRef](#)]
51. Decety, J.; Grèzes, J. Neural mechanisms subserving the perception of human actions. *Trends Cogn. Sci.* **1999**, *3*, 172–178. [[CrossRef](#)]
52. Hu, B.; Kane-Jackson, R.; Niebur, E. A proto-object based saliency model in three-dimensional space. *Vis. Res.* **2016**, *119*, 42–49. [[CrossRef](#)]
53. Silaghi, M.C.; Plänkers, R.; Boulic, R.; Fua, P.; Thalmann, D. Local and global skeleton fitting techniques for optical motion capture. In Proceedings of the International Workshop on Capture Techniques for Virtual Environments, Geneva, Switzerland, 26–27 November 1998; pp. 26–40.
54. Kurihara, K.; Hoshino, S.; Yamane, K.; Nakamura, Y. Optical motion capture system with pan-tilt camera tracking and real time data processing. In Proceedings of the 2002 IEEE International Conference on Robotics and Automation, Washington, DC, USA, 11–15 May 2002; pp. 1241–1248.
55. Zordan, V.B.; Van Der Horst, N.C. Mapping optical motion capture data to skeletal motion using a physical model. In Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation, San Diego, CA, USA, 26–27 July 2003; pp. 245–250.

56. Kirk, A.G.; O'Brien, J.F.; Forsyth, D.A. Skeletal parameter estimation from optical motion capture data. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20–25 June 2005; pp. 782–788.
57. Ghorbani, S.; Etemad, A.; Troje, N.F. Auto-labelling of Markers in Optical Motion Capture by Permutation Learning. In Proceedings of the Computer Graphics International Conference, Calgary, AB, Canada, 17–20 June 2019; pp. 167–178.
58. Vlastic, D.; Adelsberger, R.; Vannucci, G.; Barnwell, J.; Gross, M.; Matusik, W.; Popović, J. Practical motion capture in everyday surroundings. *Acm Trans. Graph.* **2007**, *26*, 35. [[CrossRef](#)]
59. Fernandez-Baena, A.; Susín Sánchez, A.; Lligadas, X. Biomechanical validation of upper-body and lower-body joint movements of kinect motion capture data for rehabilitation treatments. In Proceedings of the 2012 Fourth International Conference on Intelligent Networking and Collaborative Systems, Bucharest, Romania, 19–21 September 2012; pp. 656–661.
60. Mahmood, N.; Ghorbani, N.; Troje, N.F.; Pons-Moll, G.; Black, M.J. AMASS: Archive of motion capture as surface shapes. *arXiv* **2019**, arXiv:1904.03278.
61. Corazza, S.; Muendermann, L.; Chaudhari, A.; Demattio, T.; Cobelli, C.; Andriacchi, T.P. A markerless motion capture system to study musculoskeletal biomechanics: Visual hull and simulated annealing approach. *Ann. Biomed. Eng.* **2006**, *34*, 1019–1029. [[CrossRef](#)]
62. Mündermann, L.; Corazza, S.; Andriacchi, T.P. The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. *J. Neuroeng. Rehabil.* **2006**, *3*, 6. [[CrossRef](#)] [[PubMed](#)]
63. De Aguiar, E.; Theobalt, C.; Stoll, C.; Seidel, H.P. Marker-less deformable mesh tracking for human shape and motion capture. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8.
64. Schmitz, A.; Ye, M.; Shapiro, R.; Yang, R.; Noehren, B. Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system. *J. Biomech.* **2014**, *47*, 587–591. [[CrossRef](#)] [[PubMed](#)]
65. Giese, M.A.; Poggio, T. Morphable models for the analysis and synthesis of complex motion patterns. *Int. J. Comput. Vis.* **2000**, *38*, 59–73. [[CrossRef](#)]
66. Moeslund, T.B.; Granum, E. A survey of computer vision-based human motion capture. *Comput. Vis. Image Underst.* **2001**, *81*, 231–268. [[CrossRef](#)]
67. Grezes, J.; Fonlupt, P.; Bertenthal, B.; Delon-Martin, C.; Segebarth, C.; Decety, J. Does perception of biological motion rely on specific brain regions? *Neuroimage* **2001**, *13*, 775–785. [[CrossRef](#)] [[PubMed](#)]
68. Pollick, F.E.; Paterson, H.M.; Bruderlin, A.; Sanford, A.J. Perceiving affect from arm movement. *Cognition* **2001**, *82*, B51–B61. [[CrossRef](#)]
69. Ballan, L.; Cortelazzo, G.M. Marker-less motion capture of skinned models in a four camera set-up using optical flow and silhouettes. In Proceedings of the 3DPVT, Atlanta, GA, USA, 18–20 June 2008.
70. Rodrigues, T.B.; Catháin, C.Ó.; Devine, D.; Moran, K.; O'Connor, N.E.; Murray, N. An evaluation of a 3D multimodal marker-less motion analysis system. In Proceedings of the 10th ACM Multimedia Systems Conference, Amherst, MA, USA, 18–21 June 2019; pp. 213–221.
71. Song, Y.; Goncalves, L.; Di Bernardo, E.; Perona, P. Monocular perception of biological motion in johansson displays. *Comput. Vis. Image Underst.* **2001**, *81*, 303–327. [[CrossRef](#)]
72. Wiley, D.J.; Hahn, J.K. Interpolation synthesis of articulated figure motion. *IEEE Comput. Graph. Appl.* **1997**, *17*, 39–45. [[CrossRef](#)]
73. Grossman, E.D.; Blake, R. Brain areas active during visual perception of biological motion. *Neuron* **2002**, *35*, 1167–1175. [[CrossRef](#)]
74. Giese, M.A.; Jastorff, J.; Kourtzi, Z. Learning of the discrimination of artificial complex biological motion. *Perception* **2012**, *31*, 133–138.
75. Yi, Y.; Zheng, Z.; Lin, M. Realistic action recognition with salient foreground trajectories. *Expert Syst. Appl.* **2017**, *75*, 44–55. [[CrossRef](#)]
76. Blake, R.; Shiffrar, M. Perception of human motion. *Annu. Rev. Psychol.* **2007**, *58*, 47–73. [[CrossRef](#)] [[PubMed](#)]
77. Beintema, J.; Lappe, M. Perception of biological motion without local image motion. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5661–5663. [[CrossRef](#)] [[PubMed](#)]
78. Kilner, J.M.; Paulignan, Y.; Blakemore, S.J. An interference effect of observed biological movement on action. *Curr. Biol.* **2003**, *13*, 522–525. [[CrossRef](#)]

79. Cohen, E.H.; Singh, M. Perceived orientation of complex shape reflects graded part decomposition. *J. Vis.* **2006**, *6*, 4. [[CrossRef](#)] [[PubMed](#)]
80. Lange, J.; Georg, K.; Lappe, M. Visual perception of biological motion by form: A template-matching analysis. *J. Vis.* **2006**, *6*, 6. [[CrossRef](#)] [[PubMed](#)]
81. Gölcü, D.; Gilbert, C.D. Perceptual learning of object shape. *J. Neurosci.* **2009**, *29*, 13621–13629. [[CrossRef](#)] [[PubMed](#)]
82. McLeod, P. Preserved and Impaired Detection of Structure from Motion by “Motion-blind” Patient. *Vis. Cogn.* **1996**, *3*, 363–392. [[CrossRef](#)]
83. Daems, A.; Verfaillie, K. Viewpoint-dependent priming effects in the perception of human actions and body postures. *Vis. Cogn.* **1999**, *6*, 665–693. [[CrossRef](#)]
84. Troje, N.F.; Westhoff, C. The inversion effect in biological motion perception: Evidence for a “life detector”? *Curr. Biol.* **2006**, *16*, 821–824. [[CrossRef](#)]
85. Strasburger, H.; Rentschler, I.; Jüttner, M. Peripheral vision and pattern recognition: A review. *J. Vis.* **2011**, *11*, 13. [[CrossRef](#)] [[PubMed](#)]
86. Servos, P.; Osu, R.; Santi, A.; Kawato, M. The neural substrates of biological motion perception: An fMRI study. *Cereb. Cortex* **2002**, *12*, 772–782. [[CrossRef](#)] [[PubMed](#)]
87. Grossman, E. fMR-adaptation reveals invariant coding of biological motion on human STS. *Front. Hum. Neurosci.* **2010**, *4*, 15. [[CrossRef](#)] [[PubMed](#)]
88. Puce, A.; Perrett, D. Electrophysiology and brain imaging of biological motion. *Philos. Trans. R. Soc. Lond. B* **2003**, *358*, 435–445. [[CrossRef](#)]
89. Pyles, J.A.; Garcia, J.O.; Hoffman, D.D.; Grossman, E.D. Visual perception and neural correlates of novel ‘biological motion’. *Vis. Res.* **2007**, *47*, 2786–2797. [[CrossRef](#)]
90. Giese, M.A. Biological and body motion perception. In *Oxford Handbook of Perceptual Organization*; Oxford University Press: Oxford, UK, 2014.
91. Tlapale, É.; Doshier, B.A.; Lu, Z.L. Construction and evaluation of an integrated dynamical model of visual motion perception. *Neural Netw.* **2015**, *67*, 110–120. [[CrossRef](#)]
92. Jung, C.; Sun, T.; Gu, A. Content adaptive video denoising based on human visual perception. *J. Vis. Commun. Image Represent.* **2015**, *31*, 14–25. [[CrossRef](#)]
93. Meso, A.I.; Masson, G.S. Dynamic resolution of ambiguity during tri-stable motion perception. *Vis. Res.* **2015**, *107*, 113–123. [[CrossRef](#)]
94. Nigmatullina, Y.; Arshad, Q.; Wu, K.; Seemungal, B.; Bronstein, A.; Soto, D. How imagery changes self-motion perception. *Neuroscience* **2015**, *291*, 46–52. [[CrossRef](#)]
95. Tadin, D. Suppressive mechanisms in visual motion processing: From perception to intelligence. *Vis. Res.* **2015**, *115*, 58–70. [[CrossRef](#)]
96. Matsumoto, Y.; Takahashi, H.; Murai, T.; Takahashi, H. Visual processing and social cognition in schizophrenia: Relationships among eye movements, biological motion perception, and empathy. *Neurosci. Res.* **2015**, *90*, 95–100. [[CrossRef](#)] [[PubMed](#)]
97. Ahveninen, J.; Huang, S.; Ahlfors, S.P.; Hämäläinen, M.; Rossi, S.; Sams, M.; Jääskeläinen, I.P. Interacting parallel pathways associate sounds with visual identity in auditory cortices. *NeuroImage* **2016**, *124*, 858–868. [[CrossRef](#)] [[PubMed](#)]
98. Fu, Q.; Ma, S.; Liu, L.; Liu, J. Human Action Recognition Based on Sparse LSTM Auto-encoder and Improved 3D CNN. In Proceedings of the 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Huangshan, China, 28–30 July 2018; pp. 197–201.
99. Ji, S.; Xu, W.; Yang, M.; Yu, K. 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 221–231. [[CrossRef](#)] [[PubMed](#)]
100. Yousefi, B.; Loo, C.K. Development of biological movement recognition by interaction between active basis model and fuzzy optical flow division. *Sci. World J.* **2014**, *2014*, 238234. [[CrossRef](#)] [[PubMed](#)]
101. Yousefi, B.; Loo, C.K. Comparative study on interaction of form and motion processing streams by applying two different classifiers in mechanism for recognition of biological movement. *Sci. World J.* **2014**, *2014*, 723213. [[CrossRef](#)]
102. Yousefi, B.; Loo, C.K. Bio-Inspired Human Action Recognition using Hybrid Max-Product Neuro-Fuzzy Classifier and Quantum-Behaved PSO. *arXiv* **2015**, arXiv:1509.03789.

103. Yousefi, B.; Loo, C.K. Slow feature action prototypes effect assessment in mechanism for recognition of biological movement ventral stream. *Int. J. Bio-Inspired Comput.* **2016**, *8*, 410–424. [[CrossRef](#)]
104. He, D.; Zhou, Z.; Gan, C.; Li, F.; Liu, X.; Li, Y.; Wang, L.; Wen, S. StNet: Local and global spatial-temporal modeling for action recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 8401–8408.
105. Imtiaz, H.; Mahbub, U.; Schaefer, G.; Zhu, S.Y.; Ahad, M.A.R. Human Action Recognition based on Spectral Domain Features. *Procedia Comput. Sci.* **2015**, *60*, 430–437. [[CrossRef](#)]
106. Jhuang, H.; Serre, T.; Wolf, L.; Poggio, T. A biologically inspired system for action recognition. In Proceedings of the IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 14–21 October 2007; pp. 1–8.
107. Yamato, J.; Ohya, J.; Ishii, K. Recognizing human action in time-sequential images using hidden markov model. In Proceedings of the 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Champaign, IL, USA, 15–18 June 1992; pp. 379–385.
108. Li, Z.; Gavriluyk, K.; Gavves, E.; Jain, M.; Snoek, C.G. Videolstm convolves, attends and flows for action recognition. *Comput. Vis. Image Underst.* **2018**, *166*, 41–50. [[CrossRef](#)]
109. Wang, Y.; Mori, G. Human action recognition by semilattent topic models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 1762–1774. [[CrossRef](#)]
110. Alkurdi, L.; Busch, C.; Peer, A. Dynamic contextualization and comparison as the basis of biologically inspired action understanding. *Paladyn J. Behav. Robot.* **2018**, *9*, 19–59. [[CrossRef](#)]
111. Guo, Y.; Li, Y.; Shao, Z. DSRF: A flexible trajectory descriptor for articulated human action recognition. *Pattern Recognit.* **2018**, *76*, 137–148. [[CrossRef](#)]
112. Poppe, R. A survey on vision-based human action recognition. *Image Vis. Comput.* **2010**, *28*, 976–990. [[CrossRef](#)]
113. Fernández-Caballero, A.; Castillo, J.C.; Rodríguez-Sánchez, J.M. Human activity monitoring by local and global finite state machines. *Expert Syst. Appl.* **2012**, *39*, 6982–6993. [[CrossRef](#)]
114. Webb, B.S.; Roach, N.W.; Peirce, J.W. Masking exposes multiple global form mechanisms. *J. Vis.* **2008**, *8*, 16. [[CrossRef](#)]
115. Shu, N.; Tang, Q.; Liu, H. A bio-inspired approach modeling spiking neural networks of visual cortex for human action recognition. In Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN), Beijing, China, 6–11 July 2014; pp. 3450–3457.
116. Nweke, H.F.; Teh, Y.W.; Al-garadi, M.A.; Alo, U.R. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Syst. Appl.* **2018**, *105*, 233–261. [[CrossRef](#)]
117. Oniga, S.; Suto, J. Activity recognition in adaptive assistive systems using artificial neural networks. *Elektron. Elektrotehnika* **2016**, *22*, 68–72. [[CrossRef](#)]
118. Nguyen, T.V.; Mirza, B. Dual-layer kernel extreme learning machine for action recognition. *Neurocomputing* **2017**, *260*, 123–130. [[CrossRef](#)]
119. Layher, G.; Brosch, T.; Neumann, H. Real-time biologically inspired action recognition from key poses using a neuromorphic architecture. *Front. Neurobotics* **2017**, *11*, 13. [[CrossRef](#)]
120. Wang, L.; Ge, L.; Li, R.; Fang, Y. Three-stream CNNs for action recognition. *Pattern Recognit. Lett.* **2017**, *92*, 33–40. [[CrossRef](#)]
121. Tu, Z.; Xie, W.; Qin, Q.; Poppe, R.; Veltkamp, R.C.; Li, B.; Yuan, J. Multi-stream CNN: Learning representations based on human-related regions for action recognition. *Pattern Recognit.* **2018**, *79*, 32–43. [[CrossRef](#)]
122. Ma, M.; Marturi, N.; Li, Y.; Leonardi, A.; Stolkin, R. Region-sequence based six-stream CNN features for general and fine-grained human action recognition in videos. *Pattern Recognit.* **2018**, *76*, 506–521. [[CrossRef](#)]
123. Lu, X.; Yao, H.; Zhao, S.; Sun, X.; Zhang, S. Action recognition with multi-scale trajectory-pooled 3D convolutional descriptors. *Multimed. Tools Appl.* **2019**, *78*, 507–523. [[CrossRef](#)]
124. Kleinlein, R.; García-Faura, Á.; Luna Jiménez, C.; Montero, J.M.; Díaz-de María, F.; Fernández-Martínez, F. Predicting Image Aesthetics for Intelligent Tourism Information Systems. *Electronics* **2019**, *8*, 671. [[CrossRef](#)]
125. Wu, J.; Li, Z.; Qu, W.; Zhou, Y. One Shot Crowd Counting with Deep Scale Adaptive Neural Network. *Electronics* **2019**, *8*, 701. [[CrossRef](#)]

126. Shi, Y.; Tian, Y.; Wang, Y.; Zeng, W.; Huang, T. Learning long-term dependencies for action recognition with a biologically-inspired deep network. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 716–725.
127. Liu, C.; Freeman, W.T.; Adelson, E.H.; Weiss, Y. Human-assisted motion annotation. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.
128. Lehky, S.R.; Kiani, R.; Esteky, H.; Tanaka, K. Statistics of visual responses in primate inferotemporal cortex to object stimuli. *J. Neurophysiol.* **2011**, *106*, 1097–1117. [[CrossRef](#)] [[PubMed](#)]
129. Yue, S.; Rind, F.C. Redundant neural vision systems—Competing for collision recognition roles. *IEEE Trans. Auton. Ment. Dev.* **2013**, *5*, 173–186.
130. Mathe, S.; Sminchisescu, C. Actions in the eye: Dynamic gaze datasets and learnt saliency models for visual recognition. *IEEE Trans. Onpattern Anal. Mach. Intell.* **2015**, *37*, 1408–1424. [[CrossRef](#)]
131. Moayedi, F.; Azimifar, Z.; Boostani, R. Structured sparse representation for human action recognition. *Neurocomputing* **2015**, *161*, 38–46. [[CrossRef](#)]
132. Guha, T.; Ward, R.K. Learning sparse representations for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 1576–1588. [[CrossRef](#)]
133. Guthier, T.; Willert, V.; Schnall, A.; Kreuter, K.; Eggert, J. Non-negative sparse coding for motion extraction. In Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX, USA, 4–9 August 2013; pp. 1–8.
134. Nayak, N.M.; Roy-Chowdhury, A.K. Learning a sparse dictionary of video structure for activity modeling. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 4892–4896.
135. Dean, T.; Washington, R.; Corrado, G. Recursive sparse, spatiotemporal coding. In Proceedings of the 2009 11th IEEE International Symposium on Multimedia, San Diego, CA, USA, 14–16 December 2009; pp. 645–650.
136. Ikizler, N.; Duygulu, P. Histogram of oriented rectangles: A new pose descriptor for human action recognition. *Image Vis. Comput.* **2009**, *27*, 1515–1526. [[CrossRef](#)]
137. Guo, W.; Chen, G. Human action recognition via multi-task learning base on spatial-temporal feature. *Inf. Sci.* **2015**, *320*, 418–428. [[CrossRef](#)]
138. Shabani, A.H.; Zelek, J.S.; Clausi, D.A. Human action recognition using salient opponent-based motion features. In Proceedings of the 2010 Canadian Conference Computer and Robot Vision, Ottawa, ON, Canada, 31 May–2 June 2010; pp. 362–369.
139. Cadieu, C.F.; Olshausen, B.A. Learning intermediate-level representations of form and motion from natural movies. *Neural Comput.* **2012**, *24*, 827–866. [[CrossRef](#)] [[PubMed](#)]
140. Tian, Y.; Ruan, Q.; An, G.; Xu, W. Context and locality constrained linear coding for human action recognition. *Neurocomputing* **2015**, *167*, 359–370. [[CrossRef](#)]
141. Pitzalis, S.; Sdoia, S.; Bultrini, A.; Committeri, G.; Di Russo, F.; Fattori, P.; Galletti, C.; Galati, G. Selectivity to translational egomotion in human brain motion areas. *PLoS ONE* **2013**, *8*, e60241. [[CrossRef](#)] [[PubMed](#)]
142. Willert, V.; Toussaint, M.; Eggert, J.; Körner, E. Uncertainty optimization for robust dynamic optical flow estimation. In Proceedings of the Sixth International Conference on Machine Learning and Applications (ICMLA 2007), Cincinnati, OH, USA, 13–15 December 2007; pp. 450–457.
143. Prinz, W. Action representation: Crosstalk between semantics and pragmatics. *Neuropsychologia* **2014**, *55*, 51–56. [[CrossRef](#)]
144. Schüldt, C.; Laptev, I.; Caputo, B. Recognizing human actions: A local SVM approach. In Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK, 26 August 2004; pp. 32–36.
145. Lange, J.; Lappe, M. A model of biological motion perception from configural form cues. *J. Neurosci.* **2006**, *26*, 2894–2906. [[CrossRef](#)]
146. Willert, V.; Eggert, J. A stochastic dynamical system for optical flow estimation. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision Workshops, Kyoto, Japan, 27 September–4 October 2009; pp. 711–718.
147. Yau, J.M.; Pasupathy, A.; Fitzgerald, P.J.; Hsiao, S.S.; Connor, C.E. Analogous intermediate shape coding in vision and touch. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16457–16462. [[CrossRef](#)]
148. Escobar, M.J.; Masson, G.S.; Vieville, T.; Kornprobst, P. Action recognition using a bio-inspired feedforward spiking network. *Int. J. Comput. Vis.* **2009**, *82*, 284–301. [[CrossRef](#)]

149. Guthier, T.; Willert, V.; Eggert, J. Topological sparse learning of dynamic form patterns. *Neural Comput.* **2014**, *1*, 42–73. [[CrossRef](#)]
150. Baumann, F.; Ehlers, A.; Rosenhahn, B.; Liao, J. Recognizing human actions using novel space-time volume binary patterns. *Neurocomputing* **2016**, *173*, 54–63. [[CrossRef](#)]
151. Haghighi, H.; Abdollahi, F.; Gharibzadeh, S. Brain-inspired self-organizing modular structure to control human-like movements based on primitive motion identification. *Neurocomputing* **2016**, *173*, 1436–1442. [[CrossRef](#)]
152. Esser, S.; Merolla, P.; Arthur, J.; Cassidy, A.; Appuswamy, R.; Andreopoulos, A.; Berg, D.; McKinstry, J.; Melano, T.; Barch, D.; et al.. Convolutional networks for fast, energy-efficient neuromorphic computing. *arXiv* **2016**, arXiv:1603.08270.
153. Ward, E.J.; MacEvoy, S.P.; Epstein, R.A. Eye-centered encoding of visual space in scene-selective regions. *J. Vis.* **2010**, *10*, 6. [[CrossRef](#)] [[PubMed](#)]
154. Escobar, M.J.; Kornprobst, P. Action recognition via bio-inspired features: The richness of center-surround interaction. *Comput. Vis. Image Underst.* **2012**, *116*, 593–605. [[CrossRef](#)]
155. Goodale, M.A.; Westwood, D.A. An evolving view of duplex vision: Separate but interacting cortical pathways for perception and action. *Curr. Opin. Neurobiol.* **2004**, *14*, 203–211. [[CrossRef](#)]
156. Yousefi, B.; Yousefi, P. ABM and CNN application in ventral stream of visual system. In Proceedings of the 2015 IEEE Student Symposium in Biomedical Engineering & Sciences (ISSBES), Shah Alam, Malaysia, 4 November 2015; pp. 87–92.
157. Jellema, T.; Baker, C.; Wicker, B.; Perrett, D. Neural representation for the perception of the intentionality of actions. *Brain Cogn.* **2000**, *44*, 280–302. [[CrossRef](#)]
158. Billard, A.; Matarić, M.J. A biologically inspired robotic model for learning by imitation. In Proceedings of the fourth international conference on Autonomous agents, Barcelona, Spain, 3–7 June 2000; pp. 373–380.
159. Vaina, L.M.; Solomon, J.; Chowdhury, S.; Sinha, P.; Belliveau, J.W. Functional neuroanatomy of biological motion perception in humans. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11656–11661. [[CrossRef](#)]
160. Fleischer, F.; Caggiano, V.; Thier, P.; Giese, M.A. Physiologically inspired model for the visual recognition of transitive hand actions. *J. Neurosci.* **2013**, *33*, 6563–6580. [[CrossRef](#)]
161. Syrris, V.; Petridis, V. A lattice-based neuro-computing methodology for real-time human action recognition. *Inf. Sci.* **2011**, *181*, 1874–1887. [[CrossRef](#)]
162. Vander Wyk, B.C.; Voos, A.; Pelphrey, K.A. Action representation in the superior temporal sulcus in children and adults: An fMRI study. *Dev. Cogn. Neurosci.* **2012**, *2*, 409–416. [[CrossRef](#)]
163. Troje, N.F. Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *J. Vis.* **2002**, *2*, 371–387. [[CrossRef](#)] [[PubMed](#)]
164. Banquet, J.P.; Gaussier, P.; Quoy, M.; Revel, A.; Burnod, Y. A hierarchy of associations in hippocampo-cortical systems: Cognitive maps and navigation strategies. *Neural Comput.* **2005**, *17*, 1339–1384. [[CrossRef](#)] [[PubMed](#)]
165. Yamamoto, K.; Miura, K. Effect of motion coherence on time perception relates to perceived speed. *Vis. Res.* **2015**, *123*, 56–62. [[CrossRef](#)] [[PubMed](#)]
166. Schindler, A.; Bartels, A. Motion parallax links visual motion areas and scene regions. *NeuroImage* **2016**, *125*, 803–812. [[CrossRef](#)] [[PubMed](#)]
167. Venezia, J.H.; Fillmore, P.; Matchin, W.; Isenberg, A.L.; Hickok, G.; Fridriksson, J. Perception drives production across sensory modalities: A network for sensorimotor integration of visual speech. *NeuroImage* **2016**, *126*, 196–207. [[CrossRef](#)] [[PubMed](#)]
168. Harvey, B.M.; Dumoulin, S.O. Visual motion transforms visual space representations similarly throughout the human visual hierarchy. *NeuroImage* **2016**, *127*, 173–185. [[CrossRef](#)] [[PubMed](#)]
169. Rizzolatti, G.; Fogassi, L.; Gallese, V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2001**, *2*, 661–670. [[CrossRef](#)]
170. Breazeal, C.; Scassellati, B. Robots that imitate humans. *Trends Cogn. Sci.* **2002**, *6*, 481–487. [[CrossRef](#)]
171. Schaal, S.; Ijspeert, A.; Billard, A. Computational approaches to motor learning by imitation. *Philos. Trans. R. Soc. Lond. B* **2003**, *358*, 537–547. [[CrossRef](#)]
172. Demiris, Y.; Johnson, M. Distributed, predictive perception of actions: A biologically inspired robotics architecture for imitation and learning. *Connect. Sci.* **2003**, *15*, 231–243. [[CrossRef](#)]

173. Johnson, M.; Demiris, Y. Hierarchies of coupled inverse and forward models for abstraction in robot action planning, recognition and imitation. In Proceedings of the AISB 2005 Symposium on Imitation in Animals and Artifacts, Hatfield, UK, 12–15 April 2005; pp. 69–76.
174. Cook, J.; Saygin, A.P.; Swain, R.; Blakemore, S.J. Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. *Neuropsychologia* **2009**, *47*, 3275–3278. [[CrossRef](#)]
175. Milner, A.D.; Goodale, M.A. Two visual systems re-viewed. *Neuropsychologia* **2008**, *46*, 774–785. [[CrossRef](#)]
176. Hesse, C.; Schenk, T. Delayed action does not always require the ventral stream: A study on a patient with visual form agnosia. *Cortex* **2014**, *54*, 77–91. [[CrossRef](#)]
177. Schenk, T. No dissociation between perception and action in patient DF when haptic feedback is withdrawn. *J. Neurosci.* **2012**, *32*, 2013–2017. [[CrossRef](#)]
178. Schenk, T. Response to Milner et al.: Grasping uses vision and haptic feedback. *Trends Cogn. Sci.* **2012**, *16*, 258. [[CrossRef](#)]
179. Whitwell, R.L.; Milner, A.D.; Cavina-Pratesi, C.; Byrne, C.M.; Goodale, M.A. DF's visual brain in action: The role of tactile cues. *Neuropsychologia* **2014**, *55*, 41–50. [[CrossRef](#)]
180. Whitwell, R.L.; Milner, A.D.; Cavina-Pratesi, C.; Barat, M.; Goodale, M.A. Patient DF's visual brain in action: Visual feedforward control in visual form agnosia. *Vis. Res.* **2015**, *110*, 265–276. [[CrossRef](#)]
181. Krigolson, O.E.; Cheng, D.; Binsted, G. The role of visual processing in motor learning and control: Insights from electroencephalography. *Vis. Res.* **2015**, *110*, 277–285. [[CrossRef](#)]
182. Cavina-Pratesi, C.; Large, M.E.; Milner, A.D. Reprint of: Visual processing of words in a patient with visual form agnosia: A behavioural and fMRI study. *Cortex* **2015**, *72*, 97–114. [[CrossRef](#)]
183. Libet, B.; Wright, E.W.; Gleason, C.A. Preparation-or intention-to-act, in relation to pre-event potentials recorded at the vertex. *Electroencephalogr. Clin. Neurophysiol.* **1983**, *56*, 367–372. [[CrossRef](#)]
184. Chao, Y.W. Visual Recognition and Synthesis of Human-Object Interactions. Ph.D. Thesis, University of Michigan, Ann Arbor, MI, USA, 2019.
185. Hoshide, R.; Jandial, R. Plasticity in Motion. *Neurosurgery* **2019**, *84*, 19–20. [[CrossRef](#)]
186. Bicanski, A.; Burgess, N. A computational model of visual recognition memory via grid cells. *Curr. Biol.* **2019**, *29*, 979–990. [[CrossRef](#)]
187. Calabro, F.J.; Beardsley, S.A.; Vaina, L.M. Differential cortical activation during the perception of moving objects along different trajectories. *Exp. Brain Res.* **2019**, *2019*, 1–9. [[CrossRef](#)]
188. Grossberg, S. The resonant brain: How attentive conscious seeing regulates action sequences that interact with attentive cognitive learning, recognition, and prediction. *Atten. Percept. Psychophys.* **2019**, *2019*, 1–28. [[CrossRef](#)]
189. Wagner, D.D.; Chavez, R.S.; Broom, T.W. Decoding the neural representation of self and person knowledge with multivariate pattern analysis and data-driven approaches. *Wiley Interdiscip. Rev. Cogn. Sci.* **2019**, *10*, e1482. [[CrossRef](#)]
190. Isik, L.; Tacchetti, A.; Poggio, T. A fast, invariant representation for human action in the visual system. *J. Neurophysiol.* **2017**, *119*, 631–640. [[CrossRef](#)]

