IJIAET International Journal of Innovative Analyses and Emerging Technology

| e-ISSN: 2792-4025 | http://openaccessjournals.eu | Volume: 2 Issue: 3

Computer-Vision Based Method for Human Action Recognition

Xamidov Elnur Khamidovich, Xodjimatov Jahongir Murodovichelnur

Assistants of Fergana branch of Tashkent University of Information Technologies

Abstract: This article discusses modern approaches to recognizing human movements. The method of recognition of human movements and their application are described. The features, advantages and disadvantages of using these approaches are also described.

Keywords: touch and visual technologies, HAR, CNN, descriptors, STIP, computer vision, human action recognition.

In the modern world, human motion detection systems have become popular due to the development of sensory and visual technologies. Human Movement Detection (HAR) plays an important role in many real-world applications. EVERY goal is to recognize information about the context in which an activity takes place, including the actions of an individual or a group of people from sensors or video data. With the development of sensor and visual technologies, HAR-based systems have begun to be widely used in many real-world applications. In particular, the proliferation of small sensors has allowed smart devices to recognize human activity in context. It is also used to increase the accuracy of processes performed using video technology. Based on development methodologies and the data collection process, the approaches are divided into visual, non-visual, and mixed approaches.

The main difference between visual and other types of sensors is the method of data reception, visual sensors can present data in the form of 2D or 3D images or videos, while other sensors provide data in the form of one-dimensional signals is distinguished by the fact that Numerous reviews have been published on HAR and related processes. However, due to the large number of articles published on this topic, published reviews quickly become obsolete. For the same reason, writing a review article is an urgent task. The articles follow a review of modern methods and descriptors of human activity recognition based on neural networks.

The traditional approach. This approach to recognizing human movements is based on descriptors. This approach has become popular in the EVERY community and has achieved remarkable results in a variety of media collections. In this approach, important features are derived from a sequence of image frames and a property descriptor is created using special feature detectors. Further classification is done by training a universal support vector machine (SVM). This approach includes methods based on appearance models and spatial-time models, methods that use local binary template logic.

Spatial-temporal models. Space-time models consist of four main components: a space-time point detector (STIP), a function descriptor, a dictionary constructor, and a classifier. STIP detectors are further subdivided into dense and sparse detectors. Dense detectors such as V-FAST, Hessian detector, intensive sampling cover all video content intensively to determine the breakpoint, sparse detectors such as Harris3D Cube Detector and Implicit Shape Space Time Model (STISM) use sparse (local). Function descriptors are also divided into local and global descriptors. Local descriptors such as Cube Descriptor, Extended Accelerated Reliable Functions (ESURF) and N-jet are based on local data such as texture, color and location, while global descriptors use global data such as light change, phase change and video speed change. Dictionary builders or collection methods are based on a set of words (BOW) or a spatial state model.

ISSN 2792-4025 (online), Published under Volume: 2 Issue: 3 in March-2022

IJIAET International Journal of Innovative Analyses and Emerging Technology

| e-ISSN: 2792-4025 | http://openaccessjournals.eu | Volume: 2 Issue: 3

Local Binary Pattern (LBP). Local Binary Pattern (LBP) is a type of visual descriptor for texture classification. Since its inception, several modified versions of this description have been proposed for different classification tasks when viewing a computer. A method of detecting human movement based on LBP was proposed, along with a method of appearance change and patch adjustment. This method has been tested on various mass data sets and has proven to be effective for motion detection. Another method of detecting activity using the LBP-TOP descriptor has been proposed. In this method, the scope is divided into subfolders, and a feature histogram is created by combining the histograms of the subfolders. Using this image, they encoded the movement at three different levels: pixel level (one bit in the histogram), regional level (subfolder histogram), and global level (combination of histogram subfolders). LBP-based methods have also been used to detect human movements in many users. A multi-user human motion detection method based on contour positions and single rotation-invariant LBP and then SVM was proposed for classification.

The traditional approach. This approach to recognizing human movements is based on descriptors. This approach has become popular in the EVERY community and has achieved remarkable results in a variety of media collections. In this approach, important features are derived from a sequence of image frames and a property descriptor is created using special feature detectors. Further classification is done by training a universal support vector machine (SVM). This approach includes methods based on appearance models and spatial-time models, methods that use local binary template logic.

Spatial-temporal models. Space-time models consist of four main components: a space-time point detector (STIP), a function descriptor, a dictionary constructor, and a classifier. STIP detectors are further subdivided into dense and sparse detectors. Dense detectors such as V-FAST, Hessian detector, intensive sampling cover all video content intensively to determine the breakpoint, sparse detectors such as Harris3D Cube Detector and Implicit Shape Space Time Model (STISM) use sparse (local). Function descriptors are also divided into local and global descriptors. Local descriptors such as Cube Descriptor, Extended Accelerated Reliable Functions (ESURF) and N-jet are based on local data such as texture, color and location, while global descriptors use global data such as light change, phase change and video speed change. Dictionary builders or collection methods are based on a set of words (BOW) or a spatial state model.

Local Binary Pattern (LBP). Local Binary Pattern (LBP) is a type of visual descriptor for texture classification. Since its inception, several modified versions of this description have been proposed for different classification tasks when viewing a computer. A method of detecting human movement based on LBP was proposed, along with a method of appearance change and patch adjustment. This method has been tested on various mass data sets and has proven to be effective for motion detection. Another method of detecting activity using the LBP-TOP descriptor has been proposed. In this method, the scope is divided into subfolders, and a feature histogram is created by combining the histograms of the subfolders. Using this image, they encoded the movement at three different levels: pixel level (one bit in the histogram), regional level (subfolder histogram), and global level (combination of histogram subfolders). LBP-based methods have also been used to detect human movements in many users. A multi-user human motion detection method based on contour positions and single rotation-invariant LBP and then SVM was proposed for classification:

a) generative / uncontrolled models (e.g., Deep Belief Networks (DBN), Deep Boltzmann Machines (DBM), Limited Boltzmann Machines (RBM), and regular auto coders);

b) Discriminatory / controlled models (e.g., deep neural networks (DNN), repetitive neural networks (RNN), and convolutional neural networks (CNN));

c) hybrid models, these models use features of both models, for example Discrimination targets may be useful for generative model results. However, these models are not considered separately here.

ISSN 2792-4025 (online), Published under Volume: 2 Issue: 3 in March-2022

Methods of using uncontrolled in-depth study. Uncontrolled in-depth study models do not require target markings. In-depth study models have been studied since the 1960s, but researchers have paid little attention to these models. This was mainly due to the success of shallow models such as SVM and the lack of large amounts of data needed to train in-depth models.

The remarkable rise in the history of in-depth models was the result of the work of Hinton et al., Where a highly efficient DBN and learning algorithm were introduced, followed by feature reduction techniques. The DBN was trained step-by-step using RBM, the parameters obtained in this uncontrolled preparation phase were fine-tuned under control using feedback. Since the introduction of this effective model, there has been great interest in applying in-depth study models to a variety of applications, such as speech detection, image classification, object detection, and human movement detection. A method using video data to detect uncontrolled features was proposed to detect motion. The authors used an independent subphase analysis algorithm to study spatio-temporal properties and combined them with in-depth study methods such as separation and recognition of convolutional and motion images. RBM-trained deep faith networks (DLNs) were used to recognize human movements. The proposed method outperformed the methods with descriptors. Learning to stream video continuously without any labels is an important but difficult task. This question was solved using an uncontrolled in-depth study model. Most actions are recorded in a data-driven environment; it's hard to recognize movements from endless videos.

Uncontrolled education has played a key role in reviving research interests in in-depth study. However, it left this in the shadow of a controlled study only, as a major breakthrough in in-depth study was the use of CNNs to identify an object.

In general, it can be said that in-depth study has become a very popular field of machine learning that has outperformed traditional approaches in many computer scanning programs. A very useful feature of in-depth learning algorithms is their ability to learn properties from raw data, which eliminates the need for manual descriptors and descriptors. There are two categories of in-depth study models: uncontrolled and controlled models. DBN is a popular uncontrolled model used to recognize human movements. This model has achieved high performance in a complex data set compared to traditional hand-held collectors. CNN, on the other hand, is one of the most popular in-depth study models. Most of the existing training demonstrations apply CNNs directly to video frames or use CNN modifications to distinguish and represent spatial-temporal features. These models have also achieved remarkable results in complex data sets of human activity detection. So far, controlled in-depth study models have performed well, but some research suggests that uncontrolled learning will be much more important in the long run.

References

1. Xia L., Gori I., Aggarwal J. K., Ryoo M. S. Robot-centric activity recognition from first-person rgb-d videos // Applications of Computer Vision (WACV), IEEE Winter Conference on, 2015, pp. 357–364.
2. Soomro K., Zamir A. R. Action recognition in realistic sports videos // Springer, 2014, pp. 181–208.
3. Kihl O., Picard D., Gosselin P.-H. Local polynomial space–time descriptors for action classification // Machine Vision and Applications, 2016, vol. 27, no. 3, pp. 351–361.
4. Wang H., Schmid C. Action recognition with improved trajectories // Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 3551–3558.

ISSN 2792-4025 (online), Published under Volume: 2 Issue: 3 in March-2022

IJIAET International Journal of Innovative Analyses and Emerging Technology

| e-ISSN: 2792-4025 | http://openaccessjournals.eu | Volume: 2 Issue: 3

5. Bux A., Angelov P., Habib Z. Vision based human activity recognition: a review // Springer, 2017, pp. 341–371.
6. Кодиров, Э. С. У., & Халилов, З. Ш. (2020). Возможности и преимущества искусственного интеллекта (ИИ) и логических вычислений. Universum: техническиенауки, (6-1 (75)), 18-21.
7. Кодиров, Э. С. У., & Халилов, З. Ш. (2020). Взаимосвязи и различия между “deeplearning” и “machinelearning”. Universum: технические науки, (7-1 (76)), 23-25.
8. Okhunov, M., &Minamatov, Y. (2021). Application of Innovative Projects in Information Systems. European Journal of Life Safety and Stability (2660-9630), 11, 167-168.
9. Минаматов, Ю. (2021). УМНЫЕ УСТРОЙСТВА И ПРОЦЕССЫ В ИХ ПРАКТИЧЕСКОЙ ЭКСПЛУАТАЦИИ. Eurasian Journal of Academic Research, 1(9), 875-879.