

A Survey on Image Processing and Human Action Recognition

V.Suganya

Dept.of computer Science & Engg.

Faculty of engineering,

Avinashilingam Institute for Home Science and Higher education for Women,

Coimbatore-641 108

Abstract- Image processing is a method that takes image as input and performs some operation on it and gives image as output. It takes input as photograph or video frame and gives output as the parameters related to that image. Image processing will enhance the image by taking the useful information from it. Image processing will take the image as two dimensional signals. Two methods are used for image processing analog and digital image processing. Image processing is a growing technology and can be used in a various fields like optical imaging, medical imaging, computer vision, computer graphics etc. This paper gives a broad view of various digital image processing techniques.

Keywords – Watermarking, Haar Wavelet, DWT, PSNR

I. INTRODUCTION

Image processing is done by importing the image through optical scanner, manipulating the image by means of image enhancement and finally giving the output as the user needs. Digital image processing plays an important role in sensing remote data's. Digital image processing enhances the image by modifying its image qualities. Operation of image processing is divided into image compression, enhancement, restoration and extraction. Image taken under different conditions (poor lighting, contrast etc..) can be enhanced by applying median filter and image histogram. Image extraction can be done by dividing the image into number of segments and providing the value above the threshold value to the part of the image.

II. IMAGE PROCESSING AND VIDEO RECOGNITION TECHNIQUES

Tamer Rabie [1] introduces a new approach namely robust denoising technique for eliminating the noise outliers. This approach improves the neighbourhood estimates. The proposed technique is very helpful in denoising the noise when compared to other denoising techniques by keeping the fine details.

Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, William T. Freeman [2] implements a new approach for the blur image happened during camera shake. The Bayesian approach is used to find the blur kernel in the image. By using this blur kernel image is reconstructed using deconvolution algorithm. Moreover this approach needs less amount of user input. Author ensures that this method can be extended for advanced natural image statistics and can improve the noise level.

Joost van de Weijer and Cordelia Schmid [3] proposed a new approach called set of color descriptors that remains reliable for any changes. Color descriptor gives both the color and shape of the image features. The color descriptors when compared with many application areas, it remains unchanged for geometrical changes and decreased image quality.

Fengjun Lv and Ramakant Nevatia [4] implements a new approach for human invariant action. Author takes every human action as 2D poses and represents them as graph. Matching between the input image and 2d poses is done by Pyramid Match Kernel algorithm and best match is taken using Viterbi algorithm. By applying the Active net, performance of PMK-NUP algorithm is good. Author also says that good match between the poses is found to be good when using PMK-NUP algorithm and using Active net is very effective.

Matthew Brown and David G. Lowe [5] uses invariant local features to solve panoramic image stitching by comparing all the images. Author also uses a new approach called automatic straightening for removing wavy effect. Author also uses multi band blending algorithm for blending high frequency over the image in short range.

Dmitri Bitouk , Neeraj Kumar, Samreen Dhillon, Peter Belhumeur and Shree K. Nayar [6] uses replacement algorithm for replacing the original image. This algorithm selects the image that are similar to the input image and adjust the poses of image. This approach gives highly plausible results across a wide range of skin tones, lighting conditions and viewpoints. The author tested among 12 people and user could not identify the difference between the original image and image developed by the system.

Oncel Tuzel, Fatih Porikli and Peter Meer [7] introduced new approach for human face detection called covariance matrices as object descriptors and a novel learning algorithm on the Riemannian manifolds. This approach uses covariance features for human descriptors. The author focuses on Riemannian manifolds by giving prior information about the geometry space and this approach is used in many applications.

Saad Ali and Mubarak Shah [8] proposed a new approach for human action in videos. Author introduces kinematic features obtained from optical flow is used for representing human actions. Three dimensional information is reduced into two dimensional form by means of dominant kinematic trends or kinematic modes. Kinematic mode is used for embedding features of each video. Dynamic information is represented in terms of kinematic modes in which video classification is done by classifier algorithm.

Jingen Liu, Benjamin Kuipers and Silvio Savarese [9] suggested an idea of using attributes to recognize the human actions. Author proposed a unified approach in which attributes are manually specified. The problem for selecting the action attribute is solved by treating the attributes as latent variables and by using latent linear SVM framework. The author concluded that this approach is very critical for complex actions.

Mingtao Pei, Yunde Jia and Song-Chun Zhu[10] proposed a new algorithm for predicting the plausible actions. Event parsing algorithm is useful for both indoor and outdoor videos. Event parsing algorithm handles insertion of events and improves the atomic actions. This algorithm along with event context gives lot of information about the events which can be used to recognize the objects. Author concludes that by using this algorithm events involving multiple agents can be detected accurately.

Jiang Wang, Zicheng Liu, Ying Wu and Junsong Yuan [11] proposed a new approach for human action recognition captured by depth cameras. Depth camera produces better 3D quality image. Author introduces a new approach called local occupancy pattern in which each 3D joint is associated with the LOP feature. Author also introduces a new Fourier temporal pattern to reduce the noise. A new actionlet ensemble model is introduced in this paper which is suitable for recognizing the depth data. The author concludes that the actionlet ensemble model is best suitable for intra-class variations and for reducing the noise and errors.

Yue Gao, Jinhui Tang, Richang Hong, Shuicheng Yan, Qionghai Dai, Naiyao Zhang, and Tat-Seng Chua [12] discussed a new approach namely camera constraint-free view-based (CCFV) 3-D object retrieval algorithm in which object can be captured from any direction without any constraints. CCFV can be applied to any 3D objects and this algorithm does not have any camera constraints. CCFV can also be called as camera free constraints.

Kai Guo, Prakash Ishwar, and Janusz Konrad [13] introduces a new approach for action recognition called silhouette tunnel. In this method, video is broken into many number of segments and each segment is classified using nearest neighbor rule. Proposed method reduces complexity and is very efficient and effective.

Ming-Ming Cheng, Jonathan Warrell, Wen-Yan Lin, Shuai Zheng, Vibhav Vineet and Nigel Crook [14] implemented a approach for detecting the salient regions in images. Soft image abstraction technique is used for detecting the salient region. Soft image abstraction uses histogram quantization techniques to collect samples. This paper concludes by using hierarchical indexing mechanism high quality salient maps is obtained

Masoud Faraki, Maziar Palhang and Conrad Sanderson [15] applied Bag of Words for human action recognition. Bag of Words model uses covariance matrices as image descriptors. Author uses Log-Euclidean Bag of Words (LE-BoW) for better recognition of human action.

Keze Wang, Liang Lin, Jiangbo Lu, Chenglong Li and Keyang Shi [16] discussed about saliency detection in which foreground objects is highlighted. In saliency model similar color will be distributed in both background and foreground. Author proposes a new approach called pixelwise image saliency aggregating (PISA) used for producing fine grained image. Author uses shape-adaptive cost-volume filtering technique to achieve fine saliency value. High efficiency is obtained by using PISA approach.

Dong Li, Huiling Zhou, and Kin-Man Lam [17] proposes pore-Principal Component Analysis (PCA)-Scale Invariant Feature Transform (PPCASIFT) for face recognition. In this approach accurate face alignment is not necessary. Author uses alignment-error-insensitive and pose-invariant face verification approach for real time application. Fast and robust fitting method is used in this paper for removing outliers and to block the matching of two faces. Maximized local-matching density approach is used in which prior information on different poses is not needed.

III.CONCLUSION

This paper presents a broad view on image processing and video recognition techniques. It explains various image processing techniques used for human action recognition. This paper will be very useful for researchers to get the overview on image processing techniques.

REFERENCES

- [1] Tamer Rabie, 2005, "Robust Estimation Approach for Blind Denoising", IEEE Transactions On Image Processing, Vol. 14, No. 11, November 2005.
- [2] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis ,William T. Freeman,2006," Removing Camera Shake from a Single Photograph", ACM Transactions on Graphics (TOG),Volume 25 Issue 3,July 2006.
- [3] Joost van de Weijer and Cordelia Schmid,2006 ,"Coloring Local Feature Extraction", Book title Computer Vision – ECCV 2006Springer Berlin Heidelberg,2006
- [4] Fengjun Lv Ramakant Nevatia, "Single View Human Action Recognition using Key Pose Matching and Viterbi Path Searching", Published in Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on 17-22 June 2007.
- [5] Matthew Brown and David G. Lowe,2007," Automatic Panoramic Image Stitching using Invariant Features", International journal of computer vision, Volume 74,Issue 1,pp 59-73,2007.
- [6] Dmitri Bitouk, Neeraj Kumar, Samreen Dhillon, Peter Belhumeur and Shree K. Nayar," Face Swapping: Automatically Replacing Faces in Photographs", ACM Transactions on Graphics (TOG), Volume 27 Issue 3, August 2008.
- [7] Oncel Tuzel, Fatih Porikli and Peter Meer," Human Detection via Classification on Riemannian Manifolds",published in Computer Vision and Pattern, June 2007.
- [8] Saad Ali and Mubarak Shah,"Human Action Recognition in Videos Using Kinematic Features and Multiple InstanceLearning", IEEE Transactions On Pattern Analysis And Machine Intelligence, Volume 32 , Issue 2, December 2008.
- [9] Jingen Liu, Benjamin Kuipers and Silvio Savarese, "Recognizing Human Actions by Attributes", published in Computer Vision and Pattern, June 2011.
- [10] Mingtao Pei, Yunde Jia , Song-Chun Zhu," Parsing Video Events with Goal inference and Intent Prediction", Published in IEEE conference on computer vision(ICCV),November 2011.
- [11] Jiang Wang, Zicheng Liu, Ying Wu, Junsong Yuan," Mining Actionlet Ensemble for Action Recognition with Depth Cameras", Published in IEEE conference on computer vision and pattern, June 2012.
- [12] Yue Gao, Jinhui Tang, Richang Hong, Shuicheng Yan, Qionghai Dai, Naiyao Zhang, Tat-Seng Chua," Camera Constraint-Free View-Based 3-D Object Retrieval", IEEE Transactions On Image Processing, Vol. 21, No. 4, April 2012.
- [13] Kai Guo, Prakash Ishwar, Janusz Konrad, "Action Recognition in Video by Covariance Matching of Silhouette Tunnels", Published in IEEE conference in computer Graphics and Image Processing, October 2009.
- [14] Ming-Ming Cheng, Jonathan Warrell, Wen-Yan Lin, Shuai Zheng, Vibhav Vineet , Nigel Crook, "Efficient Salient Region Detection with Soft Image Abstraction", Published in IEEE conference on Computer Vision, December 2013.
- [15] Masoud Faraki, Maziar Palhang , Conrad Sanderson," Log-Euclidean Bag of Words for Human Action Recognition", IET Computer Vision, Volume 9, Issue 3, June 2015.
- [16] Keze Wang, Liang Lin, Jiangbo Lu, Chenglong Li , Keyang Shi, "Pisa: Pixelwise Image Saliency By Aggregating Complementary Appearance Contrast Measures With Edge-Preserving Coherence", IEEE Transactions On Image Processing, Vol. 24, No. 10, October 2015.
- [17] Dong Li, Huiling Zhou, And Kin-Man Lam, "High-Resolution Face Verification Using Pore-Scale Facial Features", IEEE Transactions On Image Processing, Vol. 24, No. 8, August 2015