Integrated Intelligent Research (IIR)
International Journal of Computing Algorithm
Volume: 05 Issue: 01 June 2016, Page No. 59- 61
ISSN: 2278-2397

A Study on Sparse Representation and Optimal Algorithms in Intelligent Computer Vision

M.Noorjahan[1] , A. Punitha[2]
[1]Research Scholar, Department of Computer Science, Bharathiyar University, Coimbatore, Tamil Nadu, India
[2]Research Supervisor, Department of Computer Applications, Bharathiyar University, Coimbatore, Tamil Nadu, India
Email: noornainar@gmail.com, apunithaganesh@yahoo.com

Abstract - Computer vision is a dynamic research field which involves analyzing, modifying, and high-level understanding of images. Its goal is to determine what is happening in front of a camera and use the facts understood to control a computer or a robot, or to provide the users with new images that are more informative or esthetical pleasing than the original camera images. It uses many advanced techniques in image representation to obtain efficiency in computation. Sparse signal representation techniques have significant impact in computer vision, where the goal is to obtain a compact high-fidelity representation of the input signal and to extract meaningful information. Segmentation and optimal parallel processing algorithms are expected to further improve the efficiency and speed up in processing.

Keywords: Computer vision, sparse techniques, computational efficiency, high-fidelity, segmentation, parallel processing.

I. INTRODUCTION

The field of Computer vision includes methods for capturing, processing, analyzing, and understanding images. High-dimensional data from the real world is used to produce either numerical or symbolic information, e.g., in the forms of decisions. The challenge of vision can be described as to build a signal-to-symbol converter like other problems in AI. These signals must be converted into symbolic representations finally. The manipulation of these representations allows the machine or organism to interact intelligently with the world.
Section II throws light on the sparse signals and representation and its efficiency in general. Section III presents two classical and central problems in computer vision namely face recognition and human pose estimation. Section IV deals with the proposed model and the pseudo algorithm. Section V presents the proposed technique i.e. parallel processing to achieve speed in the process. Section VI discusses an application where the solution of problems in computer vision like face recognition and human pose estimation can be used. Section VII concludes the study done.

II. SPARSE SIGNAL AND ITS REPRESENTATION

A sparse signal can be represented as a linear combination of relatively few base elements in a basis or an over complete dictionary. Most of compressed sensing methods harness the incoherence of dense random matrices. Gilbert and Indyk surveys a parallel line of research of using sparse matrices for recovering sparse signals. Due to its special nature, sparse representation is not sensitive to complex conditions with occlusion, noise, illumination, expression, pose and so on. Sparse coding has become extremely popular for extracting features from data as it is received, particularly when the dictionary of basis vectors is learned from unlabeled data [1].

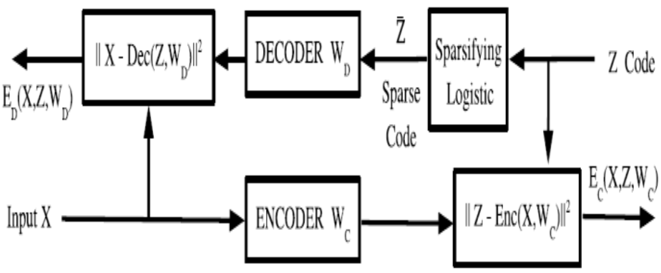

Fig. No. 1 Sparse Encoding System.

One important advantage of using sparse matrices is their efficiency in computing, in many cases allowing linear or sub linear time recovery algorithms. In recent years, researchers have changed traditional signal representation method and used the atomic over-complete system redundancy function to replace the original basis function, with the atomic elements from the library called atoms. Atomic bank selection signal system can make a good approximation of the structure, and its structure without any restrictions. To find the best linear combination with 'n' items atoms to represent a signal from atomic library is called sparse representation or highly nonlinear approximation for signal [3]. More generally, a signal x can be represented as the linear combination of T elementary waveforms, also called signal atoms, such that

$$x = \varphi\alpha = \sum_{i=1}^{T} \alpha[i]\varphi_i .$$

The following diagram shows how to obtain sparse signal by placing a non-linearity between encoder and decoder [4].

III. PROBLEMS IN COMPUTER VISION

A. Face Recognitio

Face recognition (FR) is one of the classical and central problem in computer vision. In FR the role of nose is less significant. The upper part of the face is more helpful in face recognition than the lower part [7]. Humans perform this task very easily. Sparse representation-based classification (SRC) showed the effectiveness of SR and CS for face recognition [9]. An important AI outlook on vision is that it is knowledge-driven. A long standing goal in computer vision is to understand, classify and identify human faces. The face is an expressive social organ and its image depends on age, pose, identity, viewing angle, and illumination geometry. Because of its complex nature the performance of algorithms remains very poor on these problems. Compared to other methods such as nearest subspace (NS), nearest neighbour (NN) and linear support vector machine (SVM) it has been shown that the face recognition method based sparse representation (SR) is efficient and it can obtain the best performance than other. The results were obtained by using various data base such as ORL, YALE, IMM and WIEZMANN [3]. Chaochao et al [13] mentioned that the Gaussian Face model can beat the human-level performance and can improve the accuracy to 98.52%.

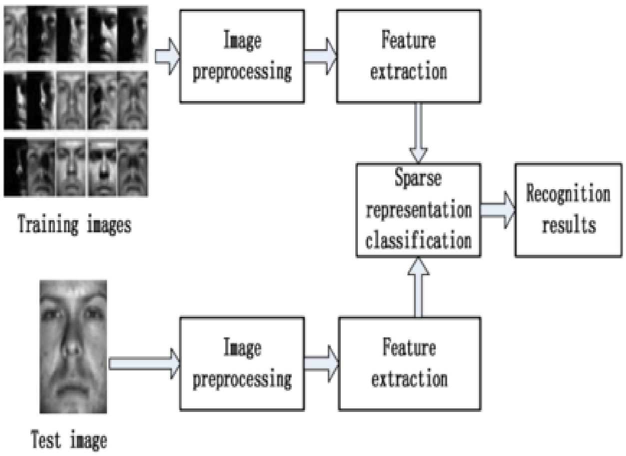

Fig. No. 2 Face recognition based on sparse representation

B. Human pose estimation:

Another important task in computer vision is Human Pose Estimation (HPE). A successful tracking system can find applications in motion capture, human computer interaction, and activity recognition. Given the initial dictionary D_o, the goal is to get a compressed dictionary D_* of size k, which encourages the signals from the same class to have very similar sparse representations. The main step in evaluating poses fit quality of a depth image from the given model. A key enabling factor is the recent development of affordable real-time depth-sensing cameras. These sensors produce the images in which each pixel has an associated depth value [5]. Dong Zhang and Mubarak Shah [10] have proposed a tree-based optimization method for human pose estimation in videos . An efficient inference algorithm based on branch-and-bound technique is proposed to solve the human pose estimation problem on loopy graphical models [11]. It has been demonstrated that both algorithms namely Body Part Classification (BPC) and Offset Joint Regression (OJR) can accurately predict the locations of body joints in 3D in super-real time from single depth or silhouette images [12].Among the many applications of computer vision such as automatic face recognition and interpretation of expression, automated medical image analysis and diagnosis, biometric-based visual identification of persons, detection of anomaly, intelligent interpretive prostheses for the blind and many, this paper focuses on the aid for the visually impaired.

IV. PROPOSED MODEL AND ALGORITHM

A. Proposed Model:

Sparse signal representation is extremely powerful tool for acquiring, representing, and compressing signals. This is mainly due to the fact that the nature of sparse representations to specify the audio and images with respect to fixed bases (i.e., Fourier, Wavelet), or as combination of such bases [8]. Efficient human pose estimation algorithms such as Branch and Bound algorithm (BB) or Particle Swarm Optimization (PSO) is used to estimate the human pose. PSO is an evolutionary computation technique frequently used for optimization tasks. Several experiments were done to evaluate the performance of PSO. Results proved that the proposed method is very promising for real-world applications which is robust [14]. It had been demonstrated that fusing algorithms such as Partial Least Squares (PLS) and Face Recognition Grand Challenge (FRGC) helped to improve performance on a difficult face-matching task [15].

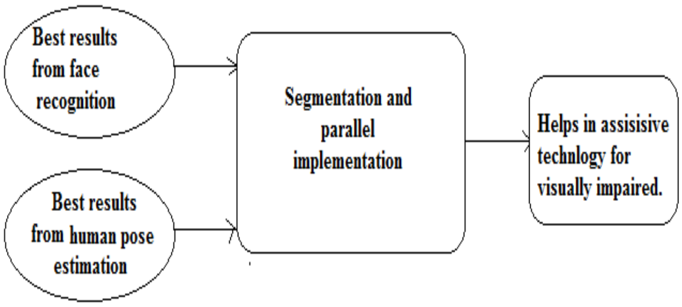

Fig. No. 3 Proposed model for assistive aid

B. Proposed Methodology:

The objective of this method is to combine the available optimal outcome of the independent existing components such as face recognition and human pose estimation. These can be inherited and implemented in the interface of the devices such as an Electronic Travel Aids (ETA) used by the Visually impaired people. As for Visually Impaired people "Ears are Eyes", the model proposed is expected to convert the image received through the camera attached to the assistive devices into voice signals. These voice signals helps to interpret the action of the people sitting around them in a conference, helps the VI persons to dynamically observe the body movement through sound signals and interact with them comfortably. This needs extensive real time operation.

C. Proposed Pseudo - algorithm

Step 1: Input the image.
Step 2: Segment the image based on efficient segmentation algorithm.
Step 3: Represent the individual segments obtained as sparse signal.
Step 4: Pass the sparse represented signal of the face segment into an optimized face recognizer.
Step 5: Pass the sparse represented signal of body and hand segment into an optimized human pose estimator (upper body).
Step 6: Combine the outcome obtained from steps 4 and 5 into voice signal in parallel.
Step 7: The output from step 6 is converted into voice signal and interpreted.

V. PARALLEL ALGORITHM

Most image processing and computer vision tasks require parallel processing since they are very computation intensive. AROB (Arrays with Reconfigurable Optical Buses) is the computational model on which the algorithms are developed. More recently, two related models namely the (AROB) array with reconfigurable optical buses and the Linear Array with a Reconfigurable Pipelined Bus System (LARPBS) have been proposed. Many algorithms have been proposed for these two relative models. These indicate that AROB is very efficient for

Volume: 05 Issue: 01 June 2016, Page No. 59- 61
ISSN: 2278-2397

parallel computation. This is possible due to the flexibility within a reconfigurable optical bus system and high bandwidth. Optical transmission can reduce the data transmission time between processors substantially. In many important applications in image processing and computer vision, the problem on hand can be reduced to a computational integer problem. Thus, it is important to develop specialized algorithms to solve these integer problems more efficiently and flexibly than with existing algorithms. Then, these basic operations help to design constant time algorithms for computer vision problems, such as histograming, shape moment computation,Hough transform, histogram modification, and median row. In relation to the product of time and the number of processors used, all of the proposed algorithms are both time and/or cost optimal [6].

VI. AN AID FOR VISUALLY IMPAIRED

Computer vision seems like a natural choice for these applications such as in replacing the lost sense of sight with an "artificial eye." A relatively large number of devices which may help the VI people possibly to replace the dog guide and the long cane altogether, have been proposed over the past 40 years. These devices termed as ETA (Electronic Travel Aids) typically utilize different types of range sensors (sonar, active triangulation systems, and stereo vision systems). Some ETA simply give an indication of the **presence** of an obstacle at a certain distance along a given direction (clear path indicators). Assistive technology is a prime example of user-centered technology. Mobile computer vision technology may be used to capture and interpret visual cues from other persons nearby, thus empowering the visually impaired user to participate more actively in the conversation. It may also help the person become aware of how he or she is perceived by others. A survey conducted in U.S with 25 visually impaired persons and two sighted specialists has highlighted some of the functionalities expected in such a system. These include understanding whether one's personal mannerisms may interfere with social interactions with others; to recognize the facial expressions of other interlocutors; and to know the names of the people nearby [2].

VII. CONCLUSION

Advances in mobile computer vision hold great promise for assistive technology. If computers can be taught to see, they may become a valuable support for those of us whose sight is compromised or lost. A variety of computer vision techniques such as face detection and human pose recognition together when implemented in parallel could help a visually impaired user correctly orient the camera and frame the picture. This is expected to be a suitable interface, which could provide these persons with a feedback mechanism so that they can participate in the conference, without feeling isolated.

References

[1] Karol Gregor and Yann LeCun, "Learning Fast Approximations of Sparse Coding". {kgregor,yann}@cs.nyu.edu Courant Institute, New York University. Communications of ACM., cacm.acm.org.

[2] Yongjiao Wang, Chuan Wang, and Lei Liang. "Sparse Representation Theory and its Application for Face Recognition"- International Journal On Smart Sensing And Intelligent Systems VOL. 8, NO. 1, MARCH 2015.

[3] Marc'Aurelio Ranzato, Christopher Poultney, Sumit Chopra, and Yann LeCun, , Courant Institute of Mathematical Sciences, New York University, New York, NY 100 {ranzato,crispy,sumit,yann}@cs.nyu.edu. " Efficient Learning of Sparse Representations with an Energy-Based Model"

[4] Matheen Siddiqui and G´erard , Human Pose Estimation from a Single View Point, Real-Time Range Sensor, IEEE August 2010.

[5] Wu, Torng, and Tsai , "Optimal Parallel Algorithms for Computer Vision Problems". National institute of standards and Technology. " Face recognition - A literature Survey".

[6] John Wright , Yi Ma, Guillermo Sapiroz , Thomas Huang, , Shuicheng Yan, Julien Mairaly. Sparse Representation For Computer Vision and Pattern Recognition. Proceedings of IEEE , March 2009.

[7] J. Wright, A. Y. Yang, A. Ganesh , S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[8] Dong Zhang, Mubarak Shah, "Human Pose Estimation in Videos", Centre for Research in Computer Vision, University of Central Florida Orlando, Florida, USA.

[9] Min Sun Murali Telaprolu Honglak Lee Silvio Savarese, "An Efficient Branch-and-Bound Algorithm for Optimal Human Pose Estimation", Department of EECS, University of Michigan, Ann Arbor, MI 48109.

[10] Jamie Shotton, Member, Ross Girshick, Andrew Fitzgibbon, Toby Sharp, Mat Cook, Mark Finocchio, Richard Moore, Pushmeet Kohli, Antonio Criminisi, Alex Kipman, Andrew Blake, Fellow, "Efficient Human Pose Estimation from Single Depth Images", CVPR 2011 special issue.

[11] Chaochao and Xiaoou, "The Face Recognition Algorithm That Finally Outperforms Humans", https://medium.com/the-physics-arxiv-blog.

[12] Luca Mussi, Spela Ivekovic , "Markerless Articulated Human Body Tracking from Multi-View Video with GPU-PSO".

[13] GPU-PSOAlice J. O'toole, Hervé Abdi, Fang JNIAG, and, P. Jonathon PHILLIPS, "Fusing Face Recognition Algorithms and Humans", IEEE transaction on systems , Man, Cybernetics, Part B 37.