International Refereed Journal of Engineering and Science (IRJES)
ISSN (Online) 2319-183X, (Print) 2319-1821
Volume 6, Issue 3 (March 2017), PP.26-30

Conceptual Framework for Anthropomorphic Simulation of Human Face for Interactive Therapeutic Telepresence Applications

Georgi Petrov[1], Polina Mihova[1], Filip Andonov[1]
New Bulgarian University, Montevideo 21 Str,

ABSTRACT. This publication is dedicated to the experimental development of conceptual framework for interactive direct marketing network based on simulated anthropomorphic agents (human faces). Aims: to discuss the main functional open system architecture and in particular openness to track and identify the mood of the user by analyzing the captured images of his face and other features. The article presents practical results, problems and possible solutions on the particular stage of development of the prototype, including project for further investigation the recognition of standardized facial expressions of emotion (anger, fear, disgust, happiness, sadness, surprise) at a perceptual level for children on the autism spectrum.
Keywords: simulated agent, prototype of human face, emotions and autism spectrum disorders (ASD).

I. INTRODUCTION

This publication discusses an open conceptual model development of distributed video monitoring system for recording and processing images of people / children through implementation of persons` anthropomorphic stylized and animated images for specific therapeutic applications [1,2]. One of the purposes of the system is to stimulate children with disabilities and special needs for perception of facial expressions, distinguishing emotions and etc.

The chosen model should allow limited scalability of hardware and to be based on an open software and hardware solutions. The current project uses hardware platform based on minicomputer module Raspberry Pi and it`s low-budget implementations Banana Pi. The operating systems of the hardware modules allow construction of network related distributed computing architecture, based on TCP/IP and multicasting applications, where synchronization functions are based on standard network protocols. The network hub is a budgetary solution by the company MikroTik, - the only one that offers integrated options for VPN, WLAN and multicasting administrative access in a single console at a low cost. The system should support network module additions that can process and analyze camera`s and microphone`s multicasting [3] incoming information and thus allow the hardware kit upgrade in order to develop 3D modeling with temporal recording and analysis movement of children/patients complex movements. [4]

1.1. Main components of the system

The basic system`s model is realized with ARM microcomputer, Linux-based solutions with open hardware and software. Author’s choice for programming language support is Python, which allows system`s performance on standard PC based devices, whish greatly expands the application areas and the future developments. The architecture choice are Raspberry Pi based solutions, but depending on the needs, objectives and available financial resources could be introduced a high-budgetarchitecture.

The basic system should have at least one 1080p@30fps resolution video camera, connected to a Raspberry Pi compatible system and a host computer. The hosting software should have options to work with standard USB cameras, but the applied Pi module significantly expands the architecture potential in terms of building multi-element video monitoring systems [5], consisting of two or more video cameras with synchronous separate video streams transmitting such as multicast messages through the local network. The primary connecting device`s protocol is an IP-based solution with specific recommendation for a wired LAN connection between the individual Pi-based cameras and the network hub. In our project, we used the MikroTik company`s product RB951G-2HnD. This type of device has number of features that successfully replace Cisco solutions whose cost exceeds ten prototypes without sacrificing the performance. The network hub can be of any type but should support a VPN, WLAN and Multicasting. Each camera is connected to the developed LAN for continuously video and audio stream broadcasting in multicasting message`s mode [6]. The selected architecture allows associating cameras and simultaneous traffic interception. The network latency is negligible due to the fact that the LAN is used locally.

At this moment, due to the Pi modules` low productivity realization of more complex algorithms for image processing at the display in real time is not possible. To preserve the interactivity of the simulated person`s image we require at least three reference images per second. Each Pi module uses open source library OpenCV or its analogue with Python interface - SimpleCV. Using a standard library allows us to scale the implementation and unify the system for monitoring and control as well as to implement it on a standard PC platform [7]. In order to work, the algorithms we use need some preliminary tasks – face detection and segmentation, head movement direction and basic mood detection. At the current stage of development, the system involves operations with one user standing near the camera module. This limits the maximum distance to the subject to 50-60cm. [8].

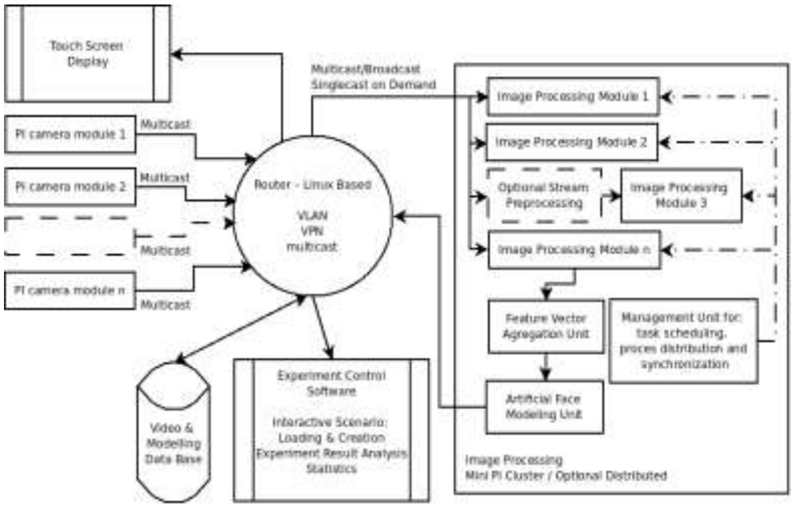

Fig.1. Basic system model

At the current stage we use a video monitor for display of anthropomorphic face image, but in the future we intent to use a mechanically controlled robot face. The cameras and microphones for patient recording aim to submit information for analysis to detect body position, person`s segmentation and facial expressions identification [9, 10]. After processing, the results generate control commands to the software for visualization of the stylized animated anthropomorphic person. The camera's stream is continuously recorded in video&modelling database. With further development it will become possible to detect basic child's body movements, hand position and response to various mechanical stimuli, etc.

II. RESOLVED ISSUES

The synchronization of the modules for segmentation and detection of child's smile is currently implemented in hardware mode, where the main control module signals alternately to the first video editing package, and when it receives valid data it sends a signal to the other. Thus, the two video editing boards allow parallel processing of twice the amount of incoming images.[5,6] Through the overclocked module Pi it is possible to achieve detection of person`s facial expressions within a second for images with 1080p@30fps resolution. It requires additional techniques for reducing image to 0.48Mpix to accelerate the processing speed of up to 2.5 frames per second and by scaling the two modules allows temporal resolution of up to 3 frames per second. For the streaming purposes is used standard application open source VLC, which supports easy creation of multicast groups and their management. [8] At the current moment, we don`t have control system to manage the process of recording the video stream.

III. HARDWARE PLATFORM LIMITATIONS

The current development does not suggest hardware acceleration through FPGA or GP-GPU modules [11, 12]. This we can't implement fast image homogenization like mode filter. Popular technique for detecting faces and human face feature extraction is based on Walsh-Hadamard Transform, including one or two preprocessing stages for color to greyscale image transform and image homogenization, like mode or median filter with recursive coefficients. All recursive algorithms for statistical image homogenization are slow to be implement on ARM based device. For this purpose we use image histogram equalization and homogenization, and multistage entropy based greyscale levels reduction as a preprocessing stage before actual face detection is done [13]. At the same time, the applied approach is universal and does not require additional knowledge for the model implementation for technical personnel with basic computer skills, such as a system administrator of a company or therapeutic center. The utilized method for processing module synchronization is hardware based. One-bit control signal is used to select the active module. In the future, it should be developed synchronization controlling system based on video frame, thus would become possible to input via network management software options for selecting, processing and segmenting video [14], which afterwards to be easily transmitted via standard TCP / IP communications control module. At the current development stage each one processing module grab frame after receiving synchronization signal from the control module. Control module should receive basic information from each processor module or cluster of modules, such as position of the person, identified emotional status and etc. In order to simulate basic emotions on the anthropomorphic face a real-time input from video processing modules is required. For smooth simulation the system should process at least 30 fps. Due to the limited number of modules (2 at the moment) the maximum achievable processed number of reference images is 3 - 4 per second. This requires the use of approximating algorithms for modeling and visualization of the animated face. Processing speed is a major bottleneck at this point.

Practical implementation

"Autism" is a generalized disorder of development that provokes much discussion, both on its etiopathogenetic mechanisms and the diverse nature of the clinical symptoms. Today therapists coalesce around two basic concept - "classic autism" that occurs from infancy, and "children with manifestations of autism spectrum", which is unlocked after 18-24 months of age. The clinical picture in both cases is dominated by the triada- problematic communication, stereotyping behavior, misunderstanding and unwillingness to accept social norms of behavior.

An important diagnostic marker in this pathology is the lack of relevance and difficulties in the perception and in the manifestation of their own emotional reactions. Difficulties with facial expressions may arise from deficits in a motivation to express positive emotions with others.

First, it should be noted that in therapeutic work with autistic main difficulty is to manage to convince, lie and provoke the child to participate in activities. ASD children are afraid of too much information and sensory satiation. Many different colors, too many elements of animation will lead to increased anxiety and demotivation and reluctance to participate in the game. Therefore pattern of elements in tasks is extremely clean both the structural elements and colors.

The study would be able to investigated the recognition of standardized facial expressions of emotion anger, fear, disgust, happiness, sadness, surprise at a perceptual level - pass a basic emotions recognition test but fail to recognize more complex stimuli involving the perception of faces or part of faces.[15]

Fig.2. Autism app [16]

IV. Discussion

Because the problem with the complexity of synchronization of individual recognition modules is serious, it is possible to combine several Raspberry Pi modules in a cluster. Such hardware solutions already exist on the market, but unfortunately the model Raspberry Pi Zero Cluster Packs is redeemed by the manufacturer. A basic requirement for the system is to combine multiple individual processing modules, in order to communicate over a LAN rather than internal highway. This allows the implementation of scalable network architecture via standard network components and protocols, where the cluster model is more convenient for uniform centralized computing tasks. Due to the flexible architecture it will be possible to include additional recognition algorithms, such as hand gesture detection. Network synchronization between different processing modules requires implementation of a complex signaling software that is not in the score of our current work.

At the moment the installation procedure is simple, but lacks the unification of the camera, processing and information aggregation modules. In order to make the system installation more scalable a centralized management application should be developed.

REFERENCES

[1]. A free interactive digital art exhibition for children, 5th - 10th July 2016, http://www.cmcplayground.com/assets/2016/06/CMC-PG-GuidevWeb.pdf [accessed Feb 23, 2017]

[2]. Philip J. Basford, Graeme M. Bragg, Jonathon S. Hare, Michael O. Jewell, Kirk Martinez, Newman R. D., R. Pau, A. Smith and Tyler Ward, Erica the Rhino: A Case Study in Using Raspberry Pi Single Board Computers for Interactive Art, Electronics 2016, 5, 35; doi:10.3390/electronics5030035, www.mdpi.com/journal/electronics (2016)

[3]. Dimitrov V .D., Analysis and trends for the use of IPTV service in global and national slace , TELECOM' 2011 13-14 October, NSTC, Sofia, BULGARIA (2011)

[4]. Berboucha M. CCTV, multiple webcams and a Raspberry Pi. Young Scientists Journal. 2015 Jan 8 [last modified: 2015 Mar 20]. Vol. 1. (2015)

[5]. Biedermann H. D., F. Dietrich, O. Handel, P. M. Kielar, M. Seitz, Using Raspberry Pi for scienic video observation of pedestrians during a music festival, arXiv:1511.00217v1 [cs.OH] (2015)

[6]. Dragomirova M, P Dulev, Developing a Learning Management Platform: A Bulgarian Case, ICEL 2104, Proceedings of the 9th International Conference on e-Learning-Proceedings of 9th International Conference on e-Learning.

[7]. Wong L., Shinji Shimojo, YuuichiTeranishi, Tomoki Yoshihisa, Jason H. Haga, Interactive Museum Exhibits with Microcontrollers: A Use-Case Scenario, PRAGMA Workshop on International Clouds for Data Science, (2015)

[8]. Fratesi A., Automated Real Time Emotion Recognition using Facial Expression Analysis, 2015, Master of Computer Science, Carleton University, Ottawa, Ontario

[9]. Mo R., A. Shaout, Portable Facial Recognition Jukebox Using Fisherfaces (Frj), International Journal of Advanced Computer Science and Applications(ijacsa), Volume 7

[10]. Swathi V., St. Fernandes, Raspberry Pi Based Human Face Detection, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 9, (2015)

[11]. Bosakova-Ardenska A.D., L.D. Bosakov, Parallel image processing with mean filter, Conference: Computer science and technologies, At Varna, Bulgaria, Available from: https://www.researchgate.net/publication/273453596_PARALLEL_IMAGE_PROCESSING_WITH_MEAN_FILTER [accessed Feb 23, 2017].

[12]. Petrov.G, F.Andonov, Development of GP-GPU supercomputers. Opportunities for applications with NVIDIA CUDA parallel processing for signals and images, NBU Telecommunication Department Yearbook, Available from: http://ebox.nbu.bg/dtk08-09/GP CUDA 2010.pdf [accessed Feb 23, 2017]

[13]. Petrov G., P. Iliev, P. Tzvetkov, Comparison of Global Histogram Methods for 2D and 3D Entropy Based Image Segmentation, 9th WSEAS International Conference on EVOLUTIONARY COMPUTING (EC'08), Sofia, Bulgaria, May 2-4, 2008

[14]. Petrov G., P. Iliev, P. Tzvetkov, Multidimensional object based image sequences segmentation, Proceedings of the 8th WSEAS Int.Conf. On ARTIFICIAL INTELLIGENCE, KNOWLEDGE ENGINEERING & DATA BASES (AIKED '09)

[15]. Boyanova, V. Stankova, M. Todorova, E., Department for the Study of communication skills in children with disorders Sofia, Slavina, 2011/85 pp./

[16]. Boyanova V., Guess the emotion, Proceedings of National conference "Interdisciplinary speech practices", 4-6 November 2016, ed.: M.Stankova, P.Mihova, ISBN 978-954-535-863-0, Vol. 2, NBU, Sofia, 2017

Polina Mihova is assistant professor in Medical Informatics and Head of Program council of Department Health and social work, NBU.
Ph.D in telemedicine and computer applications in medicine. Over 45 scientific publications; 4 textbooks, editor of 32 CDs and DVDs, author of four new academic course. Coauthor in development of 5 telemedical and medical information systems, teaching experience in medical informatics and telemedicine. Her main research interests are in information systems in biomedicine, telemedicine and terminology qualification of medical informatics, EHR for different specializations.

Filip Andonov, PhD is assistant professor in Department of Informatics at NBU. Interested in open source software in hardware, decision support systems specialist, automatic text analysis, cuber-awareness expert, human machine interaction and machine based marketing.

Georgi Petrov, PhD is head of Telecommunications department at NBU. Interested in image processing, distributed video security applications, broadband communications, multimedia streaming, FPGA's, GP-GPU, medical and genneral purpose data acquisition hardware design.