Design of an Assistant Infant Monitoring System Using Fish Robots

SEUNG Y. NA, DAEJUNG SHIN, JIN Y. KIM, MIN S. PARK, SEONG-JOON BAEK, SOWKET M. ALI, AARON PARK

Dept. of Electronics and Computer Engineering
Chonnam National University
300 Yongbong-dong, Buk-gu, Gwangju 500-757
SOUTH KOREA

Abstract: - Design of an assistant infant monitoring system requires intelligent man-machine interaction as well as ubiquitous convergence technologies since infants should be continuously monitored regardless of the location of their guardians. Moreover, the image and audio information about infants should be accessed by anyone who is interested in the states of infants. Fish robots are used for the camera platform to provide sufficient degrees of freedom of camera focusing and interesting features to the infants while preventing infants from touching the camera. To conform to the restriction of the ubiquitous environment, the amount of visual information is sufficiently reduced as far as the infants' state can be recognized. The proposed method provides effective infant monitoring results in ubiquitous environment by cellular phones, PDAs or IP devices.

Key-Words: - Fish Robot, Assistant Infant Monitoring, Ubiquitous Convergence, Man Machine Interaction

1 Introduction

In this work, we propose a general framework for assistant infant monitoring systems(AIMS) based on ubiquitous environment, and then we show the results of an exemplary implementation using a fish robot as a platform of a camera. Ubiquitous computing provides useful services to its users in their real life regardless of their location via integration of sensors, computers and their networks as described in [1]. Ideally, computing devices, electronic appliances including robots, and I/O devices such as display, sensors and actuators are networked to each other so that many kinds of applications and services are possible anytime at any place in a ubiquitous space. Furthermore, convergence of robotic technologies with ubiquitous computing gave birth to ubiquitous robots as in [2]. Mobility of robots can enhance the quality of ubiquitous computing services. Also, the horizon of ubiquitous computing applications can be broadened by the convergence of dynamics with ubiquity.

We propose a ubiquitous infant monitoring system using fish robots in an aquarium. That is, in ubiquitous environment, infants are monitored continuously by a fish robot and then the states of infants are reported to their guardians regardless of their location. Since image and audio information about infants should be accessed by anyone who is interested in the states of infants, the visual information from the camera in a fish robot as well as the audio information by a microphone is provided through wireless network to any cellular phones, PDAs or IP devices. Fish robots are chosen for the camera platform to obtain sufficient degrees of freedom of focusing since the fish robots can provide six degrees of freedom in water. Also, it is ideal for fish robots in an aquarium to prevent actively moving infants from touching the camera.

The states of infants are classified as crying, sleeping, laughing, and plain states. As an automatic attendant monitoring system, the crying state should be distinguished from the others. Since the emotional states of infants are vividly expressed in their faces, only the facial features are analyzed from the camera images. Due to the general restriction of the ubiquitous environment only the minimum amount of the visual information is processed. As soon as the crying state begins, the monitoring system should recognize it and report the state to the designated person who must get to the infants immediately and attend on them. Some features in the crying state are similar to the ones in the laughing state, particularly in low resolution images. Therefore, it is inevitable to use audio information from microphones to reduce the possibility of recognition failures. Fortunately, the features in the power spectral densities of the crying and laughing audio signals are apparently different from each other.

The ubiquitous infant monitoring system consists of a CCD camera and a microcontroller to support its camera, a fish robot to provide mobility to the camera, a microphone to get sound information, wireless communication modules and a PC server. It is important to get front face images to analyze the states of infants. But it is not easy for a fixed camera with two or three degrees of freedom to focus on an infant who is actively moving continuously. Therefore it is necessary to install the camera on a mobile robot. However, infants usually show excessive interests and touch any moving objects around them. Thus it is one of logical choices to introduce a fish robot[3] in an aquarium. A fish robot as a platform of a camera provides full degrees of freedom to make correct focusing possible. Also, it has an advantage that direct touching of the camera by infants is prevented while the fish robot provides infants with a certain degree of interest by various actions.

2 Assistant Infant Monitoring System

The physical states of infants are analyzed and classified through visual information of face images as well as the audio information from a microphone. The processed results can be accessed by networks to any cellular phones, PDAs or IP devices at any time to anyone who is interested in the states of infants. The overall structure of the monitoring system is described in Fig. 1.

There are practical restrictions to use images and sound signals in ubiquitous sensor networks due to low traffic rates, large amount of information, low capacity of microprocessors and power consumption. Therefore, images of the lowest resolution are adopted as far as facial features can be recognized. To reduce the computational burden of image processing for the microcontroller in the ubiquitous monitoring system, a small-sized camera module CMUCAM2 which has internal basic image processing units is selected. The useful features of CMUCAM2 for the monitoring system include: 1) real-time tracking of a certain color object using up to 50 frames per second; 2) a serial port of maximum 115,200 bps rate and five output ports to control motors are provided on the module.

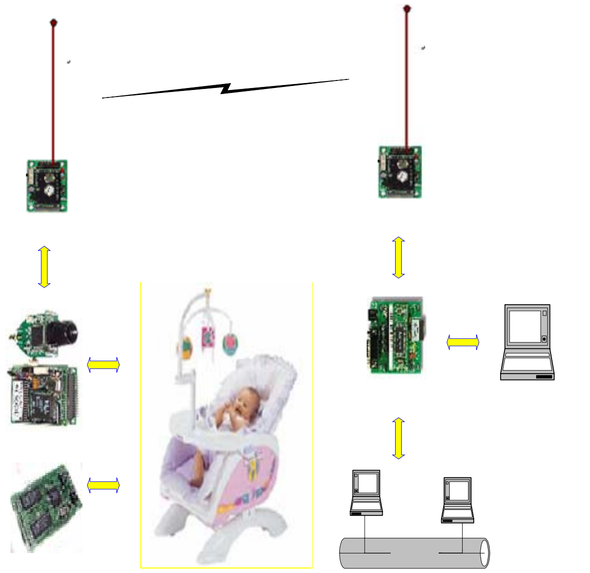


Fig. 1. Infant monitoring system

Fish robots are chosen for the camera platform to obtain sufficient degrees of freedom of focusing since the fish robots can provide six degrees of freedom in water. Also, it is ideal for fish robots in an aquarium to prevent actively moving infants from touching the camera.

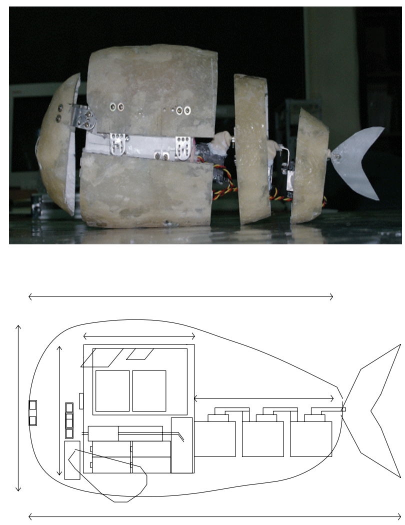


Fig. 2. A fish robot used in experiments

Although a simple small scale fish robot is considered as a platform, there are many components necessary to do path design, collision avoidance, maneuverability control and posture maintenance. A strain gauge is used to measure water pressure for depth control. MEMS-type acceleration sensors are installed to measure force changes as well as the posture of the fish body. Illumination by two LED's is for the camera images when additional light is necessary. All signals other than images are processed based on the MSP430F149 by TI. Also, user commands, sensor data and images are transmitted between robots and a host notebook PC either by Bluetooth modules or by an RF module depending on the operation depth.

Since a fish robot provides direction control using fins or propellers as in [3] for the camera to make a correct focus on infants, CMUCAM2 gets images including infant's faces. The aim of the controller of a fish robot to get an accurate focus is to settle the infant's front face at the center of images. For the tracking of the infant's face, the results of image processing are fed back to the controller of the fish robot motors. The microprocessor on the fish robot gets image data which are processed by internal routines from CMUCAM2, and then it transmits the still images to the ubiquitous home server using Bluetooth modules.

The image processing on the still images from the fish robot is done at the server PC in a U-home. The front face of an infant should be as close to the center of the final images. And then the emotional features are obtained to classify the states of an infant. The amounts of shift of the face from the center are used in next steps to the fish robot so that the motors adjust the camera focusing.

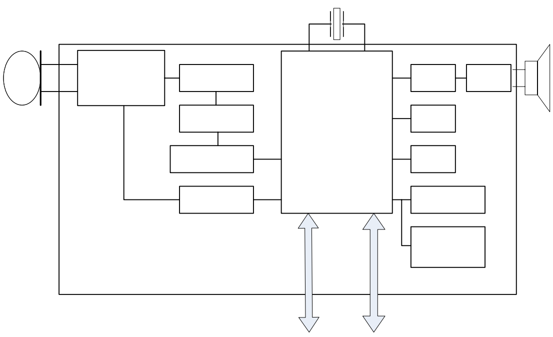


Fig. 3. Voice recognition module RSC-300

Audio information from a microphone which is fixed on a wall and connected directly to a PC is used for monitoring infants, in addition to image data. Both image and audio information are basic sources to get features for the analysis of infants' states. The final results of still images of an infant face and sound are periodically uploaded to a web server to be monitored at any time by anyone who is interested in the states of the infant.

3 Infant's State Detection

The first step of an image processing is to select only face regions out of still images transmitted from the camera module. The areas of skin color are selected based on the HIS color model. Since the emotional and physical states of infants are well expressed by the shapes of eyes and mouths, both areas are selected from face images. The overall state detection routine is shown in Fig. 4.

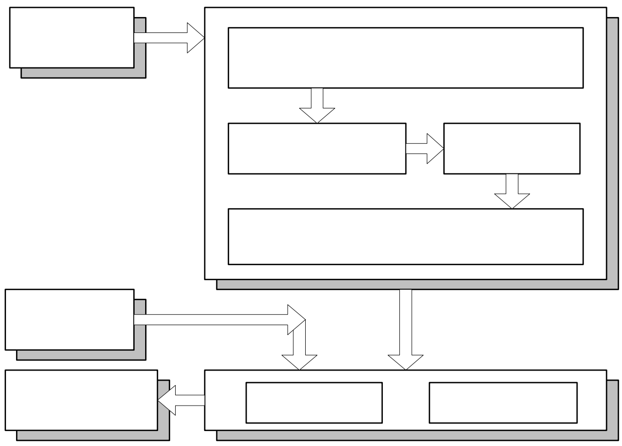


Fig. 4. The overall scheme of state detection

The shapes and sizes of eyes and mouths are the features used for the detection of crying, sleeping, laughing, and plain states. The feature vectors of eyes and mouths are shown in Fig. 5. Reduced number of feature vector elements can be used for the reduction of computation time. The horizontal line of face is obtained using the positions of two eyes. Two of the most important features from the shapes of eyes and mouths are the magnitudes and horizontal distributions of histograms. Typically the magnitudes of a mouth and eyes are small when an infant is sleeping. In the crying and laughing states both magnitudes of a mouth and eyes are relatively large. The vertical positions of the ends of both eyes indicate that they are low when in the laughing state and high in the crying state.

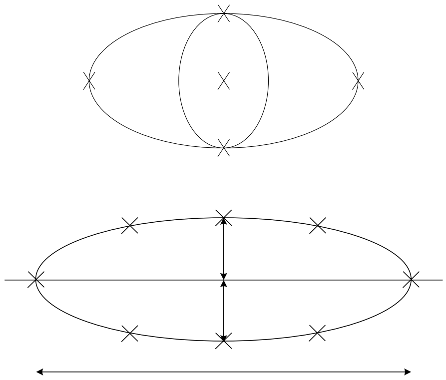


Fig. 5. The feature vectors of eyes and mouths

The most important state to be detected is the crying state. The shapes of eyes and mouths in the crying state are different obviously from those of other three states. But there are possibilities of detection failure of states particularly due to low resolution. From the viewpoint of classification by images, the crying state and the laughing state are similar to each other. Therefore, audio information from a microphone which is connected directly to the server PC, is used as well to distinguish the infant states. Each sound source is obtained in three seconds, using 16KHz, eight bits, and a single channel microphone to produce the power spectral densities of the crying and laughing states.

4 Implementation of AIMS and Results

The overall scheme of an assistant infant monitoring system(AIMS) shown in Fig. 4 is implemented based on the image and sound information in ubiquitous environment.

Since the crying state of an infant is the most important to the guardians, a certain kind of signals is sent to the designated phones immediately after the crying state is diagnosed. There are about five to ten second delays due to image capture, traffic time and image processing depending on the system. There are always possibilities of detection failure of states even though both of image and sound information are used in the monitoring system. Thus the final decision on the states of an infant should be made by the infant's guardian. At any time, anyone who cares about infants can access the still images of an infant's face through cellular phones, PDAs or IP devices(Fig. 6).

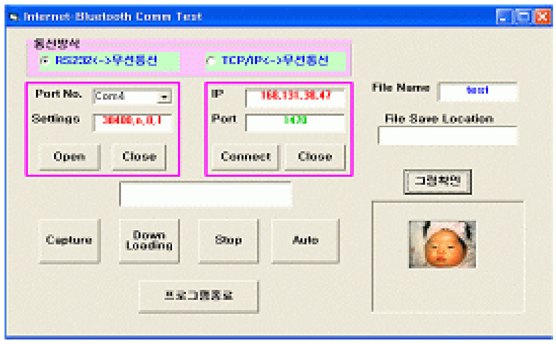


Fig. 6. The internet-based AIMS

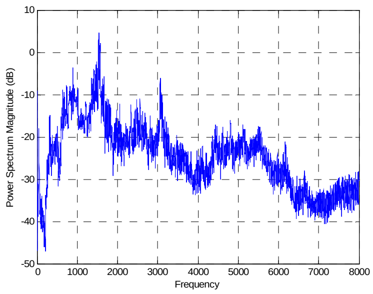
The typical experimental images of each processing step about twelve infants in crying, laughing, and sleeping states are shown below.

Fig. 7 The images in crying state

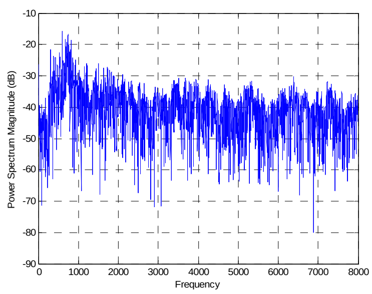
Fig. 8 The images in laughing state

Fig. 9 The images in sleeping state

For the reinforcement of classifying the crying state from the others, audio information from a microphone which is connected directly to the server PC, is used as well. The power spectral densities of the crying and laughing sound sources for typical examples are shown in Fig. 10. Each source is obtained in three seconds, using 16KHz, eight bits, and a single channel microphone.



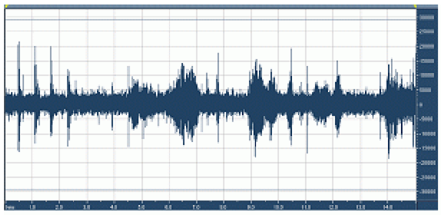
(a) crying state



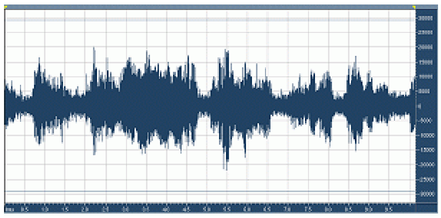
(b) laughing state

Fig. 10. Power spectral densities of captured sound sources

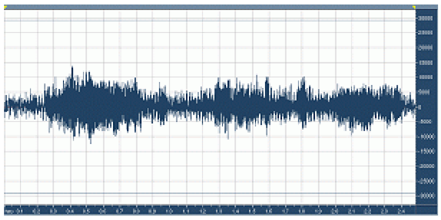
The results show that the main frequency ranges are 500Hz ~ 1,800Hz, and around 3,060Hz for crying sources and 450 ~ 850Hz for laughing sound sources. Also, there are several patterns of cry as shown in Fig. 11 by the pitch analysis.



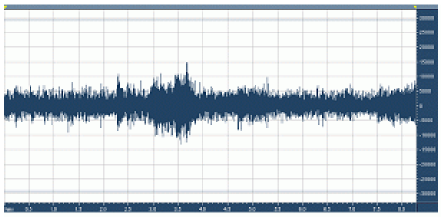
(a) intermittent crying state



(b) excessive crying state



(c) continuous crying state



(d) whimpering state

Fig. 11 The pitch analysis of cry

Any guardians can monitor the infants by mobile phones linking the server in a ubiquitous manner. Typical examples of the tested still images on the cellular phones are shown in Fig. 12.

Fig. 12. Infant monitoring by phones

5 Conclusion

An assistant infant monitoring system(AIMS) in ubiquitous environment is proposed and implemented. For the sufficient degrees of freedom for camera focusing and separation of the camera from infants, a fish robot is suggested as a camera platform. Both face images and audio information about infants are used for the diagnosis of the states of infants. The results of information processing can be accessed by anyone who is interested in the states of infants at any time and any location. There are about five to ten second delays for the final images on a server depending on the processing algorithms, pixel sizes, and specifications of a camera, communication modules and a server. Tested results prove that AIMS can provide satisfying monitoring service to anyone who cares infants.

Acknowledgements

This work was supported by the Second BK21, Chonnam National University, and RRC-HECS, CNU under grant R12-1998-007003-0.

References:

[1] S. Shepard, *RFID: Radio Frequency Identification*, McGraw-Hill, 2005

[2] Jong-Hwan Kim, "The 3rd Generation of Robotics: Ubiquitous Robot," *6th IEEE Int'l Symposium on Computational Intelligence in Robotics and Automation*, Finland, 2005.

[3] Seung Y. Na, Daejung Shin, Jin Y. Kim and Su-Il Choi, "Collision Recognition and Direction Changes Using Fuzzy Logic for Small Scale Fish Robots by Acceleration Sensor Data," *FSKD 2005*, LNAI 3614, 2005, pp. 329-338.

[4] Arati Gopalakrishnan, Sheldon Greene and Ali Sekmen, "Vision-based Mobile Robot Learning and Navigation," *RO-MAN 2005*, pp. 48-53.

[5] Wolf Kienzle, Gokhan Bakir, Matthias Franz and Bernhard Schölkopf, "Face Detection - Efficient and Rank Deficient," *18th Annual Conference on Neural Information Processing Systems*, Canada, 2004.

[6] H. Wu, Q. Chen and M. Yachida., "Face Detection from Color Images Using a Fuzzy Pattern Matching Method," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 3, 1999, pp. 557-563.

[7] C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis," *IEEE Trans. on Multimedia*, Vol. 1, No. 3, 1999, pp. 264-277.

[8] I. Kitagishi, T. Machino, A. Nakayama, S. Iwaki and M. Okudaira, "Development of Motion Data Description Language for Robots Based on eXtensible Markup Language – Realization of better Understanding and Communication via Networks," *IEEE/RSJ Int'l Conference on Intelligent Robots and Systems*, Switzerland, 2002, pp. 1145-1151.

[9] R.-J. Wai and P.-C. Chen, "Intelligent tracking control for robot manipulator including actuator dynamics via TSK-type fuzzy neural network," *IEEE Trans. Fuzzy Systems*, Vol. 12, 2004, pp. 552-560.