International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.12, (2016), pp.177-184
http://dx.doi.org/10.14257/ijsip.2016.9.12.17

Adaptive Sampling for Low Power Mobile Sign Language Video Communication

Xiaolei Chen[1], Aihua Zhang[1] and Xinzhu Yang[2]

[1]School of Electrical and Information Engineering
Lanzhou University of Technology, Lanzhou, P. R. China
[2]School of Mechanical Engineering
Lanzhou Jiaotong University, Lanzhou, P. R. China
chenxl703@gmail.com

Abstract

In this paper, we propose an adaptive sampling technique that achieves considerable energy savings while maintaining the required intelligibility level for mobile sign language video communication. The adaptive sampling scheme adjusts the sampling rate of the camera sensors dynamically according to the properties of sign language video communication and available battery power of mobile devices. Experimental results show that this adaptive scheme performs better than continuous sensing scheme in balancing the energy- intelligibility tradeoffs.

Keywords: *adaptive sampling, energy efficiency, sign language, video communication*

1. Introduction

Mobile sign language video communication (MSLVC) technique, which enables mobile devices to capture, encode, transmit, receive and decode sign language videos in real-time, can improve the efficiency of mobile communication among deaf people [1]. However, mobile devices are inherently resource constrained in terms of power, memory and bandwidth as compared to the capabilities of PCs and servers. Above all, the limited battery power is the most important factor affecting the user experience of MSLVC [2]. Moreover, the battery capacity of mobile devices does not increase at the same rate as the other components of devices, such as CPU and memory. Therefore, how to reduce the energy consumption of MSLVC without sacrificing intelligibility of sign language video has become a challenge for technological progress in MSLVC.

There have been several methods to cope with this challenge problem. In [3], sign language video is encoded and transmitted at different frame rates, when user is signing frame rate is increased to highest possible value, when user is not signing frame rate is lowered to one frame per second. In [4], the width and height of captured not signing video frames are down sampled to half before being sent to the encoder. In [5-6], an objective measure of intelligibility is included in a sign language video encoder parameter optimization by modifying a fast offline distortion-complexity algorithm, resulting in a better parameter for the rate-distortion-complexity performance. We also propose two schemes, in [7], a scheme which allocates the computational resource of sign language video encoder adaptive to available battery power and deaf people's visual system is proposed. In [8], an analytic power-rate-distortion model is proposed to obtain optimized tradeoffs among power consumption, bit rate, and distortion for MSLVC.

In this paper, we focus on the sensor sampling of mobile sign language video communication system and propose an adaptive sampling method, which can achieve the required intelligibility while conserving battery power. The sampling rate is increased or decreased to different values according to whether deaf people are singing or not and the available battery power of mobile devices. The experimental results show that this

ISSN: 2005-4254 IJSIP

adaptive sampling method can save 25% energy than continuous sensing scheme and the reduced intelligibility is less than 7%.

The remainder of this paper is organized as follows. Section 2 briefly introduces the mechanisms of sensor sampling in mobile devices. Section 3 establishes the adaptive sampling method with consideration of properties of sign language video communication and available battery power of mobile devices. Experimental results are presented in Section 4. Finally, conclusions are given in Section 5.

2. Sensor Sampling

Sensor sampling is one of the fundamental components of mobile sign language video communication system that can capture sign language information from video sensors. Camera sensors in mobile devices can run as long as required once they are turned on by users to operate. They need for an external command to finish their operations, and then to become turned off. Figure 1, shows the structure for a camera sensor operation. The first part is initialization part which is in charge of waking the sensor up and then waiting for an acknowledge response informing that the sensor is ready. The second part is called processing, this part involves sense and sleep cycles. In a sense cycle, sensors capture information continuously. In a sleep cycle, no data is captured. Sampling interval is defined as sleep interval length between two consecutive sense cycles, and sampling rate is the number of sense cycles per unit time. The third part terminates the sensor use.

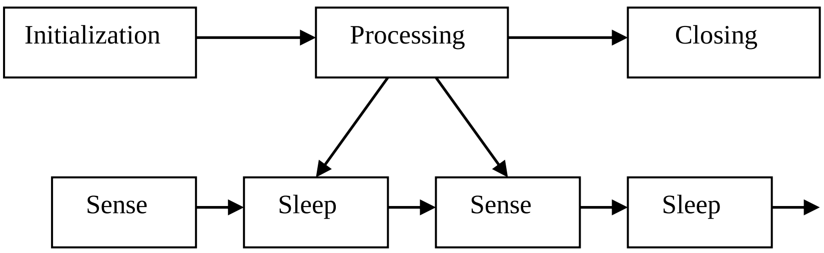


Figure 1. Sensor Sampling Process

Among these three parts, processing block can be used to achieve efficiency in battery power consumption. The longer the sense cycles, the more information will be captured, meanwhile the faster the devices' battery will be depleted. This discourages deaf people from using MSLVC because the battery may be drained during their conversation, so they have to frequently recharge their phones. On the contrary, the shorter the sense cycles, the lower the energy consumption, meanwhile the less information will be captured, and this reduces intelligibility of sign language video which also discourages deaf people from using MSLVC. Therefore, adaptive sampling rate control needs to be proposed to balance the energy- intelligibility tradeoffs.

3. Adaptive Sampling Rate Control

Sign language conversation can be divided into two parts: signing and not signing because sign language conversation involves turn-taking times when one person is signing while the other is not [3]. Signing part contains lots of sign language information, it should be captured as many as possible, not signing part is less important than signing part, so we can save battery power by lowering the sampling rate during times of not signing. Therefore, we choose properties of sign language conversation as the first factor to control sampling rate.

To include this factor into adaptive sampling rate control we must determine whether deaf people are signing or not. There are many accurate and complicated methods can be used to solve this problem. But these methods need a large number of computational resources, they are hardly implementable on battery powered mobile devices due to their high computational requirements. Therefore we propose a low complicated method to determine whether a user is signing or not.

Frames that contain a lot of fast movements in the hands and face of deaf people can result in large pixel differences, these frames should be classified as signing. Frames that contain fewer movements will result in small pixel differences, these frames can be classified as not signing.

Let $f_{k-1}(x,y)$, $f_k(x,y)$ and $f_{k+1}(x,y)$ be three successive sign language frames. We can calculate the absolute differences between frame k-1 and frame k as

$$d_{k-1,k}(x,y) = |f_k(x,y) - f_{k-1}(x,y)| \tag{1}$$

The absolute differences between frame k and frame k+1 can be calculated as

$$d_{k,k+1}(x,y) = |f_{k+1}(x,y) - f_k(x,y)| \tag{2}$$

Let $b_k(x,y)$ be the classification of the frame k, we compare $d_{k-1,k}(x,y)$, $d_{k,k+1}(x,y)$ with the pre-determined threshold *T*, if they are above the threshold, we classify the frame as signing, $b_k(x,y)$ equals to one; if they are below the threshold, we classify the frame as not signing, $b_k(x,y)$ equals to zero.

$$b_k(x,y) = \begin{cases} 1 & if \quad d_{k-1,k}(x,y), d_{k,k+1}(x,y) > T \\ 0 & else \end{cases} \tag{3}$$

We get threshold *T* by training our method on several different sign language videos.

For mobile devices, battery power decreases gradually with the working time. If battery power is sufficient, high sampling rate can be used to ensure intelligibility. If battery power is insufficient, low sampling rate must be used to reduce energy consumption, to ensure sign language video communication tasks can be finished before battery is exhausted. Therefore, we choose battery power as the second factor to determine sampling rate.

We divide battery power into three levels as Table 1, shows.

Table 1. Three Battery Power Levels

Battery power level	Sampling rate	Sign language video intelligibility
Level-1 (66%< power <=100%)	High	High
Level-2 (33%< power <=66%)	Medium	Medium
Level-3 (0%<power<=33%)	Low	Low

Level-1 targets high intelligibility and operates when more energy is available, its sampling rate is highest. Level-2 consumes less energy as compared to Level-1because its sampling rate is lower than Level-1, accordingly the sign language video intelligibility will also be reduced. Level-3 costs the least energy of all three sets by using the lowest sampling rate and its intelligibility is lowest. Figure 2, illustrates the run-time transitions from one battery power level to the other.

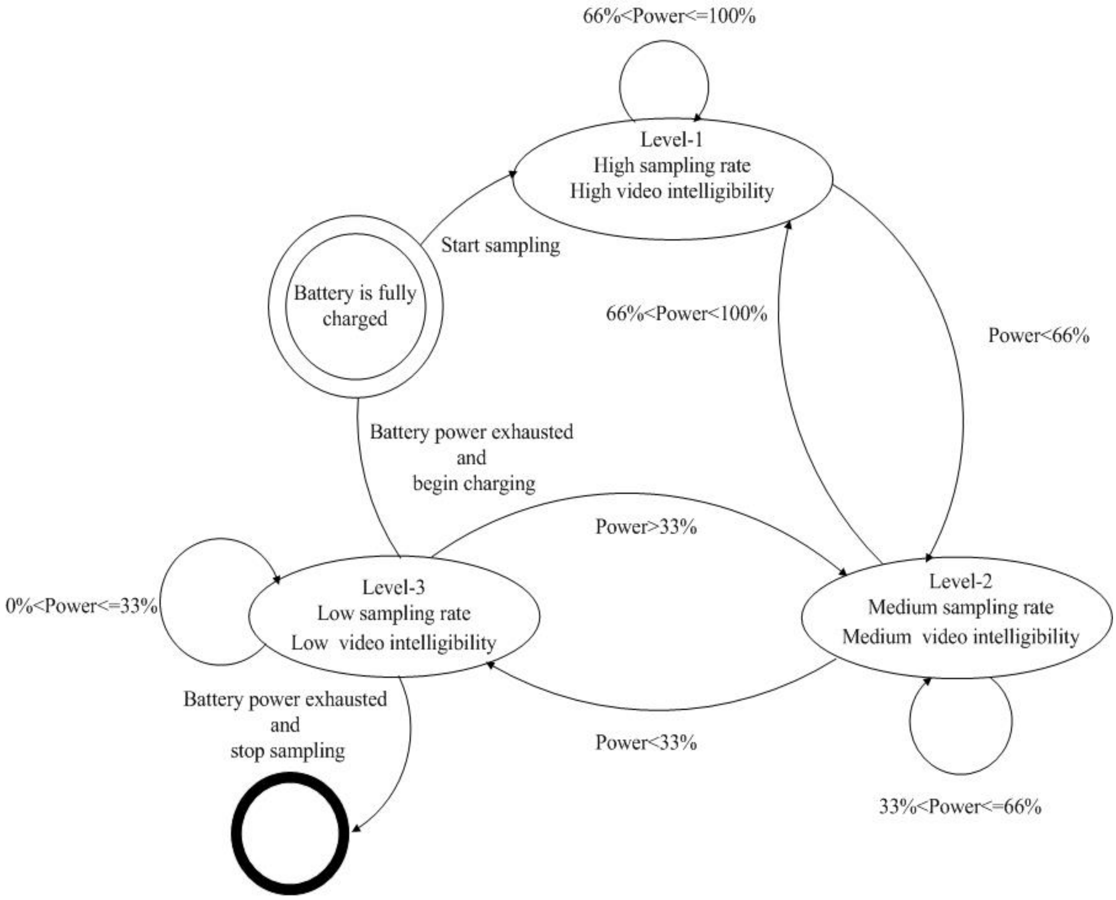


Figure 2. Flowchart of Run-Time Transitions from One Battery

Based on above discussion, with consideration of both properties of sign language video communication and battery power of mobile devices, in this paper we propose a novel adaptive sampling rate control algorithm to treat the run-time changing scenarios of available energy budgets while keeping sign language video intelligibility. The scheme is depicted in Figure 3, the major steps of the method are as follows:

Step1. Classify sign language video is signing or not and detect current battery power of mobile devices.

Step2.Sampling rate is initialized to be constant s

Step3.If user is signing and battery power is Level-1, sampling rate is increased to s^2

Step4.If user is signing and battery power is Level-2, sampling rate is increased to 2s

Step5.If user is signing and battery power is Level-3, sampling rate is increased to s+5

Step6.If user is not signing and battery power is Level-1, sampling rate is decreased to s-5

Step7.If user is not signing and battery power is Level-2, sampling rate is decreased to s/2

Step8.If user is not signing and battery power is Level-3, sampling rate is decreased to $\sqrt{s}$

Step9.If battery power is below Level-3, battery is exhausted and stops sampling

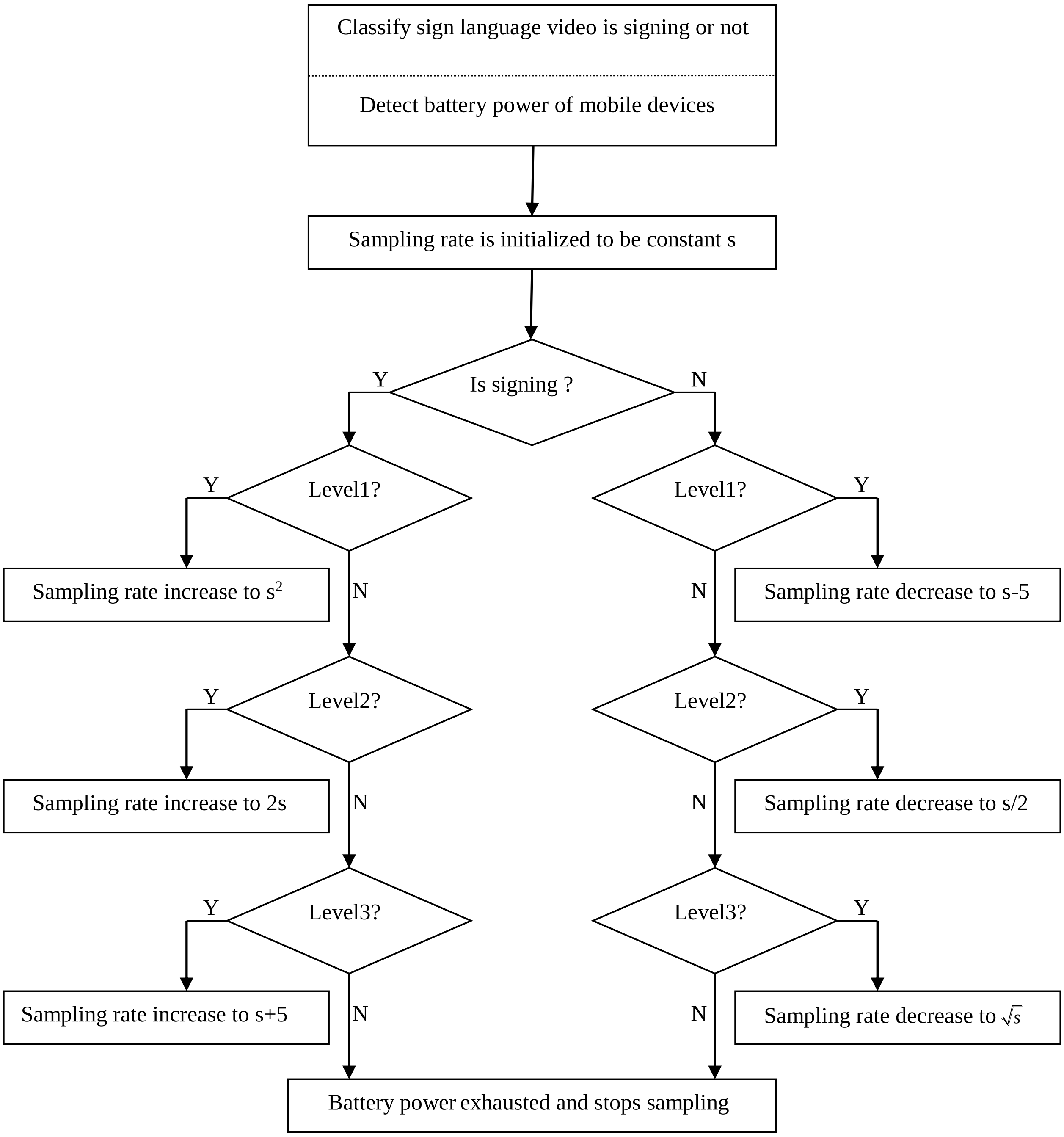


Figure 3. Flow Diagram of Adaptive Sampling Rate Control Algorithm

4. Experimental Results

In order to evaluate the proposed algorithm, we implemented it on a commercial lithium battery powered laptop equipped with camera sensor. The hardware configuration of the laptop is as follows: CPU 2.13GHz, memory 2G DDR, battery capacity 4440mAh.We also compared it with a continuous sensing scheme. Figure 4, and Figure 5, show the performance of the adaptive and continuous sensing schemes with respect to intelligibility and energy consumption.

Intelligibility is defined as the capability of a signal to be understood, given that the signal is clearly articulated, captured, transmitted, received, and perceived by the receiver, including the environmental conditions affecting these steps [9]. Figure 4, shows the comparison of intelligibility between power-adaptive and continuous sensing schemes. Intelligibility achieved by continuous scheme is 100%, whereas by power-adaptive scheme is 93%, the intelligibility drop is 7%.

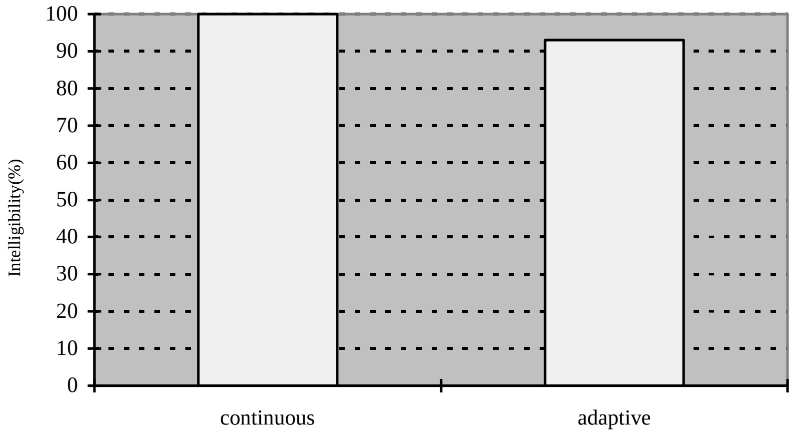


Figure 4. Intelligibility of Continuous and Adaptive Sampling Scheme

Figure 5, shows the comparison of energy usage between power-adaptive and continuous sensing schemes. Energy consumed by continuous scheme is 4000, whereas by power-adaptive is 3000, the total capacity savings is 25%.

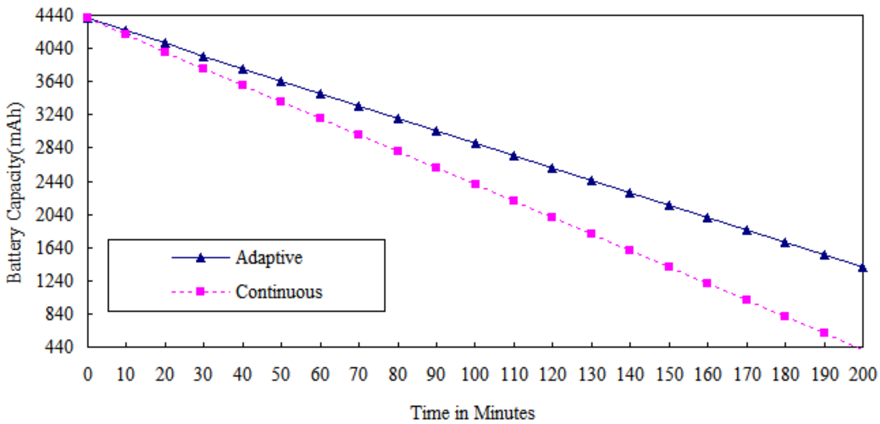


Figure 5. Battery Capacity over Time During Sign Language Video Communication

From these experiments results we can observe that the proposed power adaptive sensing scheme can achieve almost the same intelligibility and consumes 25% less energy than the continuous sensing scheme.

5. Conclusions

In this paper, we proposed a novel adaptive sampling algorithm based on properties of sign language conversation and available battery power of mobile devices. The experimental results show that the proposed adaptive sampling algorithm can save 25% battery power and the reduced intelligibility is 7%. It can be applied to meet the requirement of mobile sign language video communication under energy constraint. Furthermore, the proposed algorithm can be implemented in conjunction with other battery powered mobile multimedia applications without any impediments. For future work, the deaf people's visual system will be introduced into the current algorithm.

Acknowledgements

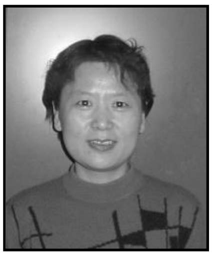
This work is supported by the National Natural Science Foundation of China under grant No. 61302116 and 61365003.

References

[1] F. M. Ciaramello and S. S. Hemami, "A computational intelligibility model for assessment and compression of American sign language video", IEEE Transactions on Image Processing, vol. 20, **(2011)**, pp. 3014-3028.
[2] M. Satyanarayanan, P. Bahl, R. Caceres and N. Davies, "The case for VM-based cloudlets in mobile computing", IEEE Pervasive Computing, vol. 8, no. 4, **(2009)**, pp. 14-23.
[3] N. Cherniavsky, A. Cavender, E. Riskin and R. Ladner, "Variable frame rate for low power mobile sign language communication", Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility, Tempe, USA, **(2007)** October 15-17.
[4] J. Tran, T. Johnson, J. Kim, R. Rodriguez, S. Yin, E. Riskin, R. Ladner and J. Wobbrock, "A web-based user survey for evaluating power saving strategies for seaf users of MobileASL", Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility, Orlando, USA, **(2010)** October 25-27.
[5] F. Ciaramello, R. Vanam, J. Chon, S. Hemami, E. Riskin and R. Ladner, "Rate-distortion-complexity optimization of an H.264/AVC encoder for real-time videoconferencing on a mobile device", Proceedings of 5th International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, USA, **(2010)** January 14-16.
[6] R. Vanam, F. Ciaramello, E. Riskin, S. Hemami and R. Ladner, "Joint rate-intelligibility-complexity optimization of an H.264 video encoder for American sign language", Proceedings of Western New York Image Processing Workshop, Rochester, USA, **(2009)** September 10-12.
[7] X. Chen, A. Zhang and X. Yang, "Power-aware regions-of-interest computational resource allocation for mobile sign language video encoding", International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 6, no. 6, **(2013)**, pp. 1-12.
[8] X Chen, A Zhang, D. Lin and X. Yang, "Power-rate-distortion optimization for mobile sign language video encoding", ICIC Express Letters, vol. 8, no. 6, **(2014)**, pp. 1697-1703.
[9] J. Tran, R. Rodriguez, E. Riskin and J. Wobbrock, "A web-based intelligibility evaluation of sign language video transmitted at low frame rates and bitrates", Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, Washington, USA, **(2013)** October 21-23.

Authors

Xiaolei Chen, Male, was born in Henan, China in 1979. He received his Ph.D. degrees from Lanzhou University of Technology in 2014. His research interests include image, video processing and wireless communication.



Aihua Zhang, Female, was born in Hebei, China in 1964, She received her Ph.D. degrees from Xi'an Jiaotong University in 2005. Since 2004, she has been with School of Electrical and Information Engineering at Lanzhou University of Technology as a Full Professor. Her main research interests are biomedical signal processing.

Xinzhu Yang, Female, was born is Hunan, China in 1981, She took the Bachelor's degree and the Master Science in Electronic Engineering at the Lanzhou Jiaotong University in 2003 and 2010, respectively. Her main research interests are signal processing and education management.