International Journal of Control and Automation
Vol.7, No.4 (2014), pp.285-294
http://dx.doi.org/10.14257/ijca.2014.7.4.25

Reduction in Encoding Redundancy over Visual Sensor Networks

Hyungkeuk Lee[1], Hyunwoo Lee[1], Won Ryu[1] and Kyounghee Lee[2]

[1]Electronics and Telecommunications Research Institute
[2]Pai Chai University

{hkeuklee, hwlee, wlyu}@etri.re.kr, leekhe@pcu.ac.kr

Abstract

Visual sensor networks (VSN) are wireless sensor networks in which each sensor has video capture and processing capability.Power consumption may be examined for encoding, transmitting, and receiving subsystems, and research has been performed on minimizing these power levels in parallel. When multiple camera modules of a visual sensor node are aimed at the same objects with different fields of view (FOVs), the captured images may overlap. Such overlapped FOVs give rise to encoding redundancy over the VSN and also lead to increased power consumption among adjacent nodes. The power-rate-distortion (P-R-D) is determined and used to construct an optimization problem for minimizing power consumption of each node, hence maximizing node lifetime. The optimal solution provides distributed power allocation and node scheduling over the VSN at the same time, via simple information sharing, resulting in network lifetime maximization. .

Keywords: *visual sensor networks, visual sensor node, overlapped field of view (FOV), power consumption, power-rate-distortion, residual energy*

1. Introduction

Visual sensor networks (VSNs) have significant differences with respect to other wireless sensor networks (WSNs), owing to the relatively large volume of data sensed by VSNs and the associated processing overhead. Each visual sensor node in a VSN is able to process image or video data locally using one or more cameras, and to extract detailed information for a wide range of applications such as video surveillance, emergency response, environmental tracking, and health monitoring.

Network lifetime maximization for conventional WSNs has been extensively discussed in the literature. The methods described in [1, 2], however, have limited applicability to VSNs, which require accounting for the power consumption required to process image or video data. Therefore, to deal with this power management issue, it is necessary to consider not only the transmission power but also the compression power, *i.e.*, the encoding power.In [3], the concept of accumulative visual information was introduced as a means for measuring the amount of visual information collected by a VSN. The minimization of video distortion in VSNs via optimizing power allocation in VSNs was investigated, while He et al. investigated the resource-distortion optimization problem for video encoding and transmission over VSNs in [4]. In order to achieve optimal power management for visual sensor nodes, it is necessary to determine rates of power consumption for the three main modules: the encoder, transmitter, and receiver. Thus the design of system protocols for VSNs needs to take into account the power consumption of the encoding module. In general, the higher the compression applied by an encoder algorithm (resulting in lower compressed rates for a given quality), the higher computation overhead becomes and the more processing power is required. Hence, for a

ISSN: 2005-4297 IJCA

given level of source distortion, the processing power can be modeled as a decreasing function of the compressed bit rate.

In this paper, to optimize these power consumption factors, we focus on an overlapped field of view (FOV), *i.e.*, an overlapped region captured by multiple cameras. Thus, gathering of redundant information from the overlapped FOVs can control overhead by reducing the amount of data routed through the network. Besides, those cooperative methods assumed that high correlation occurs when the two cameras capture an object or objects regardless of the viewing directions. Since the correlation should be measured through visual processing in the unit of macroblock (MB) or pixel and can be significantly changed according to the viewing direction, we detect the overlapped region by using the correlation in the unit of MB.

2. Proposed System Descriptions

We assume that the VSN consists of multiple visual sensor nodes and a sink node and that each sensor node can be controlled remotely by a user. Multiple nodes may acquire visual information independently of each other; moreover, each node has its own battery (i.e., power supply), and the remaining energy level of each node may be monitored individually. Each sensor node makes an effort to extend its lifetime and to conserve energy by cooperating with adjacent nodes as much possible. Table 1shows the notation used in this paper.

2.1. Power Consumption Model

Table 1. Description of notations

	Description
i	the index of visual sensor node
$\boldsymbol{N}$	the set of visual sensor nodes, $\forall i \in \boldsymbol{N}$
n	the MB index
$\boldsymbol{N}_i$	the MB index set of node$i$, $\forall i \in \boldsymbol{N}_i$
P_n	the encoding power for the n^{th} MB
$P_{s,i}$	the encoding power for node i
$P_{t,i}$	the transmit power for node i
$P_{r,i}$	the receive power for node i
P_i	the total power consumption for node i, $P_i = P_{s,i} + P_{t,i} + P_{r,i}$
P_i^{res}	the residual energy for node i
P_i^{tot}	the total energy constraint for node i
D_n	the distortion of the n^{th} MB after source coding
D_i	the total distortion of node i
R_n	the bitrate for the n^{th} MB

For node i $(i \in \boldsymbol{N})$, the total energy constraint P_i^{tot} is the sum of the residual energy, source-encoding power, data-transmit power, and data-receive power expended at the current time interval and at previous time intervals:

$$P_i^{tot} = P_i^{res}(t) + P_{s,i}(t) + P_{t,i}(t) + P_{r,i}(t) + \sum_{t'=1}^{t-1}\left(P_{s,i}(t') + P_{t,i}(t') + P_{r,i}(t')\right)$$

$$= P_i^{res}(t) + P_i(t) + \sum_{t'=1}^{t-1} P_i(t'),$$

where t is the time index and $P_i(t) = P_{s,i}(t) + P_{t,i}(t) + P_{r,i}(t)$. For brevity, we omit the time index from equations throughout this paper.

In general, network lifetimerelies on the minimum lifetime among the nodes over the network [1]. In this paper, the lifetime of node i is given by

$$T_i = P_i^{tot}/P_i, \ \forall i \in \boldsymbol{N}, \ (1)$$

while the network lifetime is determined by the minimum node lifetime:

$$T_{net} = \min_{i \in \boldsymbol{N}} P_i^{tot}/P_i \,.$$

However, under the assumption that the network lifetime can be refreshed at every time interval, the definition of the lifetime (1) can be represented in terms of the residual energy at the current time, rather than the initial total energy:

$$T_i = P_i^{res}/P_i, \qquad \forall i \in \boldsymbol{N},$$

The network lifetime can then be expressed as a function of the residual energy and the power consumption.

Assume that each visual sensor node is equipped with a discrete cosine transform (DCT)-based encoder such as H.264/AVC [5], the DC component of the DCT coefficients can been modeled using a Gaussian distribution, and the AC component as a Laplacian[6]. Muller [7] showed that a generalized Gaussian distribution (GGD). Without loss of generality, the transform coefficients can be modeled by GGD, and the rate-distortion function is represented in closed form [8]:

$$D_n = \sigma_n^2 \cdot 2^{-\gamma R_n g(P_n)},$$

where σ_n^2 is the source variance of the n^{th} MB, γ is a constant for controlling model accuracy, and $g(\cdot)$ is a function representing the power consumption model in [9][10]. The total distortion for node i is then

$$D_i = \sum_{n \in \boldsymbol{N}_i} D_n = \sum_{n \in \boldsymbol{N}_i} \sigma_n^2 \cdot 2^{-\gamma R_n g(P_n)}. \qquad (2)$$

Assume $g(\cdot) = (\cdot)^{1/\alpha}$, (2) becomes

$$D_n = \sigma_n^2 \cdot 2^{-\gamma R_n P_n^{1/\alpha}}.$$

It is assumed that the video sequence is encoded at a constant quality D, which is a target distortion at the encoder.The encoding power $P_{s,i}$ is given by

$$P_{s,i} = \sum_{n \in \boldsymbol{N}_i} P_n = \sum_{n \in \boldsymbol{N}_i} \left(\frac{1}{\gamma_n R_n} \log \frac{\sigma_n^2}{D} \right)^{\alpha} = \sum_{n \in \boldsymbol{N}_i} \frac{a_n}{R_n^{\alpha}}. \qquad (3)$$

2.2. Power Minimization Problem for Visual Sensor Node

The power minimization problem for source encoding can be expressed as

$$\min_{\mathrm{R_n}} \quad P_{s,i} = \sum_{n \in \boldsymbol{N}_i} P_n = \sum_{n \in \boldsymbol{N}_i} \frac{a_n}{R_n^{\alpha}}$$

$$\text{subject to} \quad \sum_{\forall n \in \boldsymbol{N}_i} R_n \le R_i,$$

where R_i is the bit rate constraint. To determine the optimum rate R_n and power P_n for each MB, a *Lagrangian* multiplier approach can be employed [10]. The optimal rate and the minimal power consumption are then given by

$$R_n = \frac{(\alpha \cdot a_n)^{1/\alpha+1}}{\sum_{\forall n \in \boldsymbol{N}_i} (\alpha \cdot a_n)^{1/\alpha+1}} R_i,$$

$$P_{s,i} = \frac{A_i}{R_i^{\alpha}} \sum_{n \in \boldsymbol{N}_i} a_n^{\frac{1}{\alpha+1}}, \qquad A_i = \left[\sum_{\forall\, m \in \boldsymbol{N}_i} a_m^{\frac{1}{\alpha+1}}\right]^{\alpha}. \tag{4}$$

In [11], the transmission power was modeled by

$$P_{t,i}(R_i) = C_i(d_i) \cdot R_i = \left(c_0 + c_1 d_i^{c_2}\right) \cdot R_i, \tag{5}$$

where C_i is the transmit consumption function, which is a function of the distance d_i; c_0 is the energy cost of the transmission electronics; c_1 is a coefficient corresponding to the energy cost of the transmission amplifier; and c_2 is the path-loss exponent [12]. Since the power consumption for receiving is lower than for transmitting or for encoding, the optimal rate R_i^* can be obtained using a power tradeoff between $P_{s,i}$ and $P_{t,i}$ by

$$R_i^* = \arg\ \min_{R_i} P_{s,i}(R_i) + P_{t,i}(R_i). \tag{6}$$

The optimal rate in (6) can be solved by using the Karush-Kuhn-Tucker (KKT) conditions, and the optimal power consumption can be obtained using (4) and (5).

3. Power Consumption Optimization for Multi-Visual Sensor Nodes

By removing visual redundancy from overlapped FOVs, the encoding rate can be reduced and the associated processing power can be conserved with the effect of extending the network lifetime. For longer network lifetime, the node with more power should process the FOV.

3.1. Power Minimization among Visual Sensor Nodes

Assume that the overlapped FOV can be figured out, visual sensor nodes can conduct a cooperative strategy for minimization of joint power consumption. To this end, the encoding power of node i can be rewritten as

$$P_{s,i} = \sum_{\forall n \in \boldsymbol{N}_i} P_n = \sum_{\forall n \in \boldsymbol{N}_i - \boldsymbol{N}_i \cap \boldsymbol{N}_j} P_n + \sum_{\forall n \in \boldsymbol{N}_i \cap \boldsymbol{N}_j} P_n$$

where $\boldsymbol{N}_i \cap \boldsymbol{N}_j$ is the index set corresponding to the overlapped FOV between nodes i and j. We now construct two disjoint subsets in the overlapped FOV:

$$\boldsymbol{N}_i \cap \boldsymbol{N}_j = \widetilde{\boldsymbol{N}}_i \cup \widetilde{\boldsymbol{N}}_j, \quad \widetilde{\boldsymbol{N}}_i \cap \widetilde{\boldsymbol{N}}_j = \emptyset$$

where $\widetilde{\boldsymbol{N}}_i$ is one set of MBs, encoded at node i, and $\widetilde{\boldsymbol{N}}_j$ is the other set of MBs, encoded at node j. Dividing the overlapped FOV into the encoding regions $\widetilde{\boldsymbol{N}}_i$ and $\widetilde{\boldsymbol{N}}_j$ is a key step in reducing the visual sensor nodes' power consumption. The encoding power consumption for node i after dividing the overlapped FOV is thus given by

$$P_{s,i} = \sum_{\forall n \in \boldsymbol{N}_i - \boldsymbol{N}_i \cap \boldsymbol{N}_j} P_n + \sum_{\forall n \in \widetilde{\boldsymbol{N}}_i} P_n = \frac{A_i}{R_i^{\alpha}} \left(\sum_{\forall n \in \boldsymbol{N}_i - \boldsymbol{N}_i \cap \boldsymbol{N}_j} (a_n)^{\frac{1}{\alpha+1}} + \sum_{\forall\, n \in \widetilde{\boldsymbol{N}}_i} (a_n)^{\frac{1}{\alpha+1}} \right)$$

and$P_{s,j}$is obtained similarly.

The next step for optimization is to compose the two optimal sets $\widetilde{N}_i^*$ and $\widetilde{N}_j^*$ so as to obtain the least total power consumption over the whole visual sensor network. The cooperative optimization problem for two nodes (as the number of nodes may increase, it is easily expanded) can be represented by

$$\min_{R_i,R_j} P_i + P_j = \left(P_{s,i} + P_{t,i}\right) + \left(P_{s,j} + P_{t,j}\right)$$

$$\text{subject to} \quad \bar{a}_{i,j} = \sum_{\forall n\in \widetilde{\boldsymbol{N}}_i} (a_n)^{\frac{1}{\alpha+1}} + \sum_{\forall m\in \widetilde{\boldsymbol{N}}_j} (a_m)^{\frac{1}{\alpha+1}} ,$$

where $\bar{a}_{i,j} = \sum_{\forall n\in \boldsymbol{N}_i \cap \boldsymbol{N}_j} (a_n)^{1/(\alpha+1)}$. This constraint means that the sum of $a_n^{\frac{1}{\alpha+a}}$ corresponding to the overlapped FOV is a constant. Using some manipulations, the optimization problem can be reformulated as

$$\min_{R_i,R_j} \frac{A_i}{R_i^{\alpha}}(B_i + \bar{a}_i) + C_i \cdot R_i + \frac{A_j}{R_j^{\alpha}}\left(B_j + \bar{a}_j\right) + C_j \cdot R_j$$

$$\text{subject to} \quad \bar{a}_{i,j} = \bar{a}_i + \bar{a}_j ,$$

where $B_i = \sum_{\forall n\in \boldsymbol{N}_i - \boldsymbol{N}_i \cap \boldsymbol{N}_j} (a_n)^{1/(\alpha+1)}$, $B_j = \sum_{\forall m\in \boldsymbol{N}_j - \boldsymbol{N}_i \cap \boldsymbol{N}_j} (a_m)^{1/(\alpha+1)}$, $\bar{a}_i = \sum_{\forall n\in \widetilde{\boldsymbol{N}}_i} (a_n)^{1/(\alpha+1)}$ and $\bar{a}_j = \sum_{\forall m\in \widetilde{\boldsymbol{N}}_j} (a_m)^{1/(\alpha+1)}$, respectively. Note that $\bar{a}_i$ and $\bar{a}_j$ represent the exclusive encoding portion for nodes $i$and j.

3.2. Lifetime Maximization of Visual Sensor Nodes

Here, we accomplish power minimization based on residual energy, formulated as an optimization problem. The network lifetime is defined as follows:

$$\max T_{net} = \max \min_{i\in \boldsymbol{N}} \frac{P_i^{res}}{P_i} , \qquad \forall i \in \boldsymbol{N} .$$

To solve the above problem, we introduce the *Lagrangian* relaxation method

$$G(P_i, \kappa_i) = \sum_{\forall i\in \boldsymbol{N}} \ln \frac{P_i}{P_i^{res}} + \sum_{\forall i\in \boldsymbol{N}} \{\kappa_i (P_i^{res} - P_i)\},$$

whereκ_i is the *Lagrangian* multiplier and P_i is a function of R_i. The complementary slackness condition of this problem is then satisfied:

$$\frac{\partial G}{\partial R_i} = \frac{1}{P_i(R_i)} \frac{\partial P_i(R_i)}{\partial R_i} = 0.$$

4. Simulation Results

In this section, we use the system parameters employed for simulations:

$$D = 32.5 \text{ (dB)}, \gamma = 55.54 (\mathrm{W}^{3/2}/\mathrm{Mb/s}), \alpha = 2.5, c_0 = 0.5 \text{ (J/Mb)},$$
$$c_1 = 1.3 * 10^{-8} \left(\mathrm{J/Mb/m^4}\right) and\ c_2 = 4.$$

Figure 1 represents the power-rate-distortion relation described in Eq. (3) for a given distortion D. As the rate increases in 'Conventional One,' the power consumption decreases up to a certain level and increases after the rate point. At the point where power consumption is minimized, we can find the optimal rate and power values. The optimal power level was found to be lower by0.11×10^5 Watt than in the conventional case. We use two sensor nodes as shown in Figure 2. The bold line indicates the overlapped FOV.

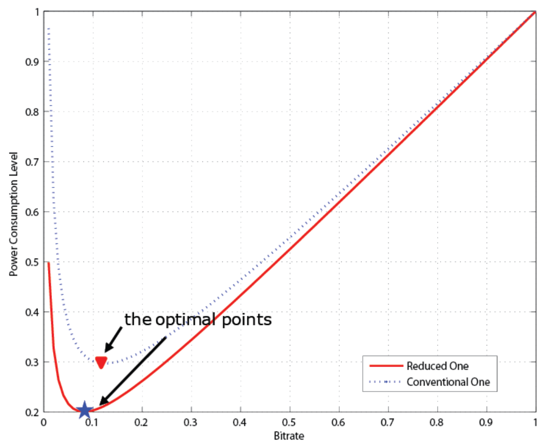

Figure 1. Normalized power level according to the normalized rate

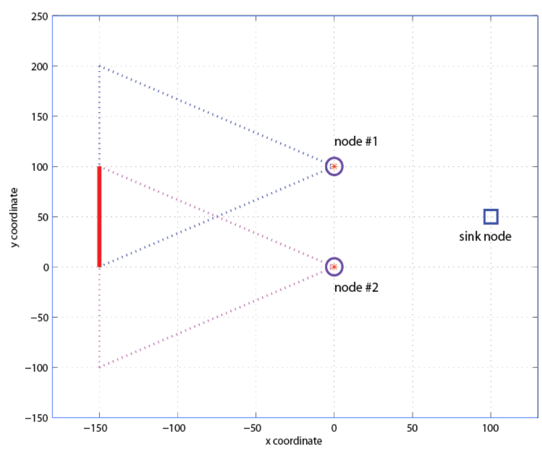

Figure 2. Placement of visual sensor nodes and the sink node

Figure 3 represents the power consumption level corresponding to the three optimization methods. In our configuration, the total energy of node 1 is 80% of that of node 2. The power consumption for 'Cooperation' is lower than that for 'No Consideration' due to the encodingregion reduction. Moreover, in 'Consideration withLifetime', the power consumption depends on the residual energy; hence node 1 consumes lesspower than node 2 because of the greater residual energy of node 2.

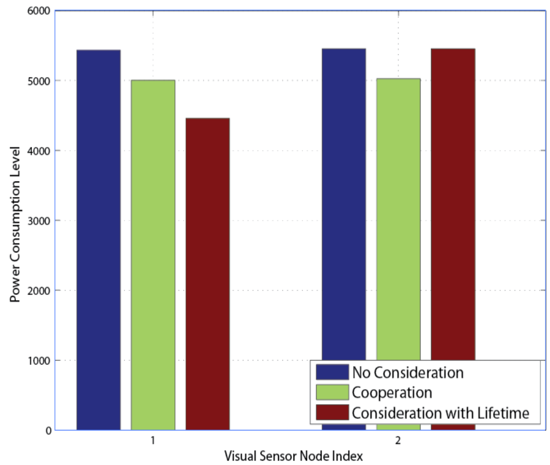

Figure 3. Power consumption level according to three optimization methods

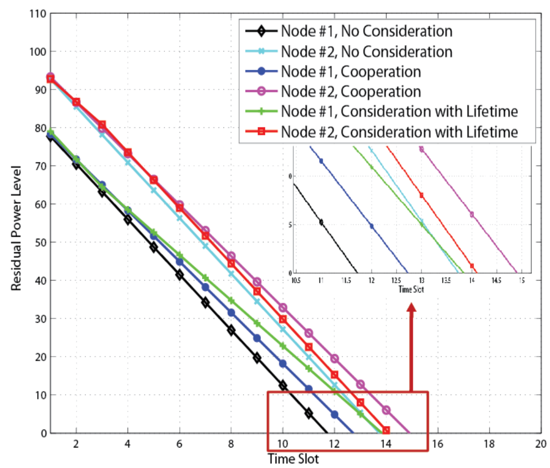

Figure 4. Relationship between the lifetime and the residual energy

Figure 4 shows the results of these methods in terms of network lifetime. The cross-points on the x-axis represent lifetimes according to the power management method. As shown in Figure 4, the network lifetime for 'No Consideration' is the least, while the network lifetime for'Consideration with Lifetime' extends over the widest range. Another discovery is that the lifetimes of nodes 1 & 2 under 'Consideration with Lifetime' approach the same value.

5. Conclusion

This paper dealt with the problem of network lifetime maximization via minimizing the total power consumption of each visual sensor node. We investigated power optimization for joint processing and transmission of captured data based on the P-R-D relation. The optimization was designed to minimize encoding power while reducing encoding redundancy for the overlapped FOVs, and to minimize transmit power while reducing the number of transmitted bits. Each node needs to share some information with other nodes to exploit this optimization in a distributed fashion, and the optimization initiated each node ultimately leads to prolongation of the network lifetime.

Acknowledgements

This research was supported by the MSIP(Ministry of Science, ICT & Future Planning), Korea, under the R&D program supervised by the KCA(Korea Communications Agency) [KCA-2013-12-912-03-001].

References

[1] J. H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks", IEEE/ACM Trans. on Networking, vol. 12, no. 4, **(2004)** August, pp. 609 – 619.
[2] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in wireless sensor networks", IEEE Trans.WirelessCommun., vol. 5, no. 8, **(2006)** August, pp. 2185 – 2193.
[3] Z. He and D. Wu, "Accumulative visual information in wireless video sensor network: definition and analysis", in IEEE ICC, **(2005)** May; Seoul, Korea,
[4] Z. He and C. W. Chen, "From rate-distortion analysis to resource-distortion analysis", in IEEE WIRELESSCOM, **(2005)** June; Maui.
[5] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard", vol. 17, no. 9, **(2007),** pp. 1103 – 1120.
[6] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images", IEEE Trans. Image Processing, vol. 9, no. 10, **(2000)** October, pp. 1661 – 1666.
[7] F. M¨uller, "Distortion shape of two-dimensional DCT coefficients of natural images", Electronics Letters, vol. 29, no. 22, **(1993)** October, pp. 1935 – 1936.
[8] T. Berger, Rate Distortion Theory, Prentice-Hall, Englewood Cliffs, NJ, **(1984)**.
[9] Z. He, W. Cheng and X. Chen, "Energy minimization of portable video communication devices based on power-ratedistortion optimization", vol. 18, no. 5, **(2008)** May, pp. 596 – 608.
[10] Y. He, I. Lee and L. Guan, "Distributed algorithms for network lifetime maximization in wireless visual sensor networks", IEEE Trans. Circuits Syst. Video Tech., vol. 19, no. 5, **(2009)** May, pp. 704 – 718.
[11] W. B. Heinzelman, A. P. Chandrakasan and H. Balakrishnan, "An application-specific protocol acchitecture for wireless microsensor networks", IEEE Trans. Wireless Commun., vol. 1, no. 4, **(2002)** October, pp. 660 – 670.
[12] T. S. Rappaport, "Wireless Communications: Principles and Practice", NJ: Prentice Hall, Englewood Cliffs, 2nd edition, **(2002)**.

Authors

Hyungkeuk Lee

He received the B.S., M.S., and Ph.D.degrees in electrical and electronic engineering fromYonsei University, Seoul, Korea, in 2005, 2006, and2011, respectively. By 2013, he was with Samsung Electronics, Inc.,Suwon, Korea. He is currently with ETRI, Daejeon, Korea. In 2008, he was a Visiting Researcherwith the Laboratory for Image and Video Engineering,University of Texas at Austin, Austin, involvedin research under guidance of Prof. A. C. Bovik. Hiscurrent research interests include wireless resourceallocation based on economics, video coding, cross-layeroptimization, image/video quality assessments, compressive sensing, 3-Dimage processing, and computer vision.

Hyunwoo Lee

He received M.S. and Ph.D. degrees in 1995 and 2005, respectively, in Korea Aerospace University (KAU). He is currently a principal research engineer and team leader in cloud convergence service networking research team, smart screen convergence research department, ETRI. His main research interests include heterogeneous wireless access network, Mobile P2P, open IPTV platform in NGN. His current research interests include cloud computing and platform.

Won Ryu

He received the BS degree in computer science and statistics from Pusan National University, Busan, South Korea, in 1983, and the MS degree in computer science and statistics from Seoul National University, Seoul, South Korea, in 1987. He received his PhD degree in information engineering from Sungkyunkwan University, Kyonggi, South Korea, in 2000. Since 1989, he has been a managing director with the Smart screen convergence research department, ETRI, Daejeon, Korea. Currently, his research interests are IPTV, Smart TV, IMT-advanced, and convergence services and networks and etc.

Kyounghee Lee

He received his M.S. and Ph.D. degrees in Korea Advanced Institute of Science and Technology (KAIST) in 2000 and 2006, respectively. From 2006 to 2013, he worked for Electronics and Telecommunications Research Institute (ETRI). Since 2013, he has been a faculty member of Department of Computer Engineering in Pai Chai University. He has expertise in international standardization on mobility management technology area as a Rapporteur of Q.22/SG13 in ITU-T. His research interests include mobile communications, Internet QoS, real-time multimedia service and cloud computing.