일반논문 (Regular Paper)
방송공학회논문지 제24권 제6호, 2019년 11월 (JBE Vol. 24, No. 6, November 2019)
https://doi.org/10.5909/JBE.2019.24.6.1085
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

초협대역 전송 시스템상에서 H.265/HEVC 부호화 저해상도 비디오에 대한 주관적 화질 평가

에이. 에프. 엠. 샤합 우딘[a], 엠에스티. 시라잠 모니라[a], 정 태 충[a], 김 동 현[b], 최 증 원[b], 전 기 남[c], 배 성 호[a]

Subjective Video Quality Evaluation of H.265/HEVC Encoded Low Resolution Videos for Ultra-Low Band Transmission System

A. F. M. Shahab Uddin[a], Mst. Sirazam Monira[a], TaeChoong Chung[a], Donghyun Kim[b], Jeung Won Choi[b], Ki Nam Jun[c], and Sung-Ho Bae[a]

요 약

본 논문에서는 저해상도 감시(surveillance) 비디오에 대한 주관적 화질 평가를 수행한다. 본 논문은 저해상도 영상에 맞는 초협대역 전송 환경을 고려하여 비디오 압축을 수행한 다음, 압축된 비디오의 주관적 화질 성능을 측정하였다. 데이터의 일반성을 확보하기 위해 다양한 시/공간 복잡도를 가지는 감시 비디오를 수집하였다. 저해상도 감시 비디오 압축은 H.265/HEVC로 수행되었으며 3개의 목표 비트(20, 45, 65 kbps)에 대해 압축을 수행했다. 주관적 화질 평가 결과, 수집한 대부분의 저해상도 감시 비디오는 45 kbps이상으로 압축될 경우 주관적 화질 열화가 거의 발생하지 않는 것으로 나타났다. 뿐만 아니라, 기존 개발된 객관적 영상 화질 측정 방법을 이용해 예측된 화질과의 상관관계를 비교하는 실험을 진행했고, 실험 결과 현존하는 대부분의 객관적 화질 측정 방법이 초협대역 전송 조건에서 저해상도 감시 비디오의 화질을 제대로 예측하지 못하는 것을 확인하였다. 이는 초협대역 전송 기반 저해상도 감시 비디오에 대한 새로운 객관적 영상 화질측정 기법이 개발되어야 함을 시사한다.

Abstract

In this paper, we perform a subjective quality assessment on low-resolution surveillance videos, which are encoded with a very low target bit-rate to use in an ultra-low band transmission system and investigate the encoding effects on the perceived video quality. The test videos are collected based on their spatial and temporal characteristics which affect the perceived quality. H.265/HEVC encoder is used to prepare the impaired sequences for three target bit-rates 20, 45, and 65 kbps and subjective quality assessment is conducted to evaluate the quality from a viewing distance of 3H. The experimental results show that the quality of encoded videos, even at target bit-rate of 45 kbps can satisfy the users. Also we compare objective image/video quality assessment methods on the proposed dataset to measure their correlation with subjective scores. The experimental results show that the existing methods poorly performed, that indicates the need for a better quality assessment method.

Keywords : Subjective Video Quality Evaluation, HEVC, Ultra-Low Band Transmission System

Ⅰ. Introduction

Recently, a massive number of video surveillance cameras have been installed over the world for security purpose which stores and/or transmit an extensive amount of video data, and increases the bandwidth utilization considerably that is the most essential aspect in any communication system. As a result, to transmit real-time surveillance video data over an ultra-low bandwidth transmission system, efficient video encoding techniques are indispensable. Therefore, it is crucially important to study the effects of video encoding techniques on the perceived quality of surveillance video data. This paper carefully creates a low-resolution surveillance video database and investigates the effects of a very low bite-rate encoding system on the perceived quality, by performing a subjective quality evaluation. Further, we compare the objective quality assessment methods on the collected database to observe their prediction accuracy. First, we collect 54 reference video sequences including RGB and IR videos, by considering their spatial and temporal complexity and then apply H.265/HEVC [8] video encoding techniques with three target bit-rates which results with 162 impaired video sequences. Then we collect subjective ratings via double stimulus impairment scale (DSIS) method recommended by ITU-R BT500-11 [9] and measure the mean opinion score (MOS).

The remainder of the paper is organized as follows. Sec. II presents the systematic way of selecting the proper test sequences. Sec. III describes the subjective evaluation process. Sec. IV compares the objective image and video quality assessment methods. Sec. V concludes the paper.

a) 경희대학교(Department of Computer Science and Engineering, Kyung Hee University)
b) 국방과학연구소(Agency for Defense Development)
c) LIG넥스원(LIG Nex1)
‡ Corresponding Author : 배성호(Sung-Ho Bae)
E-mail: shbae@khu.ac.kr
Tel: +82-31-201-2593
ORCID: https://orcid.org/0000-0003-2677-3186
※ This work has been supported by Agency for Defense Development (ADD).(UD180024ED).
・ Manuscript received July 17, 2019; Revised September 6, 2019; Accepted September 6, 2019.

Ⅱ. Systematic Selection of Test Sequences

To perform a subjective quality assessment, first we need to collect the test sequences in an appropriate way. The candidate video sequences that are used in this paper, are collected from two sources: “Standard Video Sequences” provided by high efficiency video coding (HEVC) [1] and MPEG [2]; and Collected Video Sequences. The standard video sequences are of RGB color format, while the Collected video sequences includes videos of RGB and IR color format. Based on camera motion, all of the sequences are divided into two groups: (i) stationary camera sequences and (ii) moving camera sequences. As a result, we have total six categories of video sequences, i.e., Standard RGB Moving Camera, Standards RGB Stationary Camera, Collected RGB Moving Camera, Collected RGB Stationary Camera, Collected IR Moving Camera and Collected IR Stationary Camera sequences. Figure 1 shows the categories of candidate video sequences. All of the candidate sequences are of YUV420 format with 8-bit color depth and 10 seconds long. The spatial resolution and frame rate for moving sequences are specified to (QCIF) and 10 fps respectively. On the other hand, the spatial resolution and frame rate for stationary sequences are specified to (CIF) and 5 fps, respectively. Table 1 shows the specifications of the candidate sequences at a glance. According to ITU-R BT.1210 [3], an appropriate set of test materials should have various characteristics, because the perceived quality of video sequences under test depends on their spatial and temporal characteristics. Consequently, in this paper we classified the candidate video sequences based on the spatial complexity and temporal variation, to select a proper set of test sequences. To examine the spatial complexity and temporal variation, we have used the MPEG-7

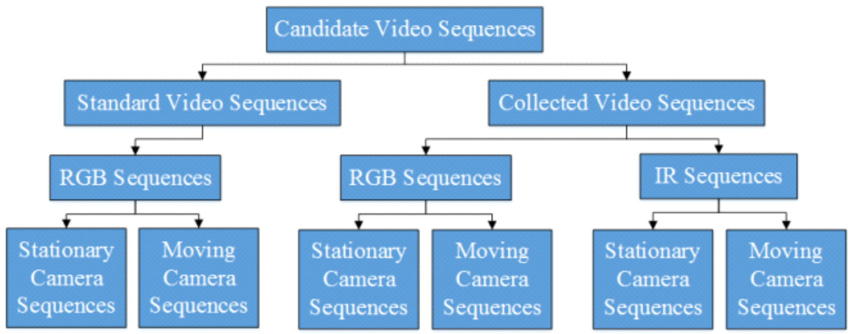

그림 1. 후보 비디오 시퀀스 카테고리
Fig. 1. Categories of candidate video sequences

표 1. 후보 시퀀스 특징
Table 1. Specifications of the candidate sequences

Source	Color format	Type	Resolution	Frame rate (fps)	Duration (second)	# of Frames	# of Sequences
Standard	RGB	Moving camera	(QCIF)	10	10 s	100	32
		Stationary camera	(CIF)	5	10 s	50	26
Collected	RGB	Moving camera	(QCIF)	10	10 s	100	64
		Stationary camera	(CIF)	5	10 s	50	68
	IR	Moving camera	(QCIF)	10	10 s	100	38
		Stationary camera	(CIF)	5	10 s	50	53
						Total	281

Edge Histogram Descriptor [4] and the average magnitude of motion vectors [3], respectively. Among several motion estimation algorithms, we have used block motion vectors to compute the temporal variation. We denote the spatial and temporal variation as and respectively. For each of the six categories, we select nine test sequences which result with a total 54 test sequences. The selection process for a single category is described as follows: (i) we first collect the specified number of candidate sequences (as shown in Table 1) of various spatial and temporal complexities; (ii) we then classify the candidate sequences into nine different spatial-temporal classes based on the and values. The spatial-temporal classes include LL, ML, HL, LM, MM, HM, LH, MH, HH, where the first letters indicate Low (L), Medium (M) and High (H) spatial complexity and second letters indicate the temporal complexity in the same manner; (iii) from each spatial-temporal classes, one sequence is chosen as a test sequence. Thereby nine test sequences

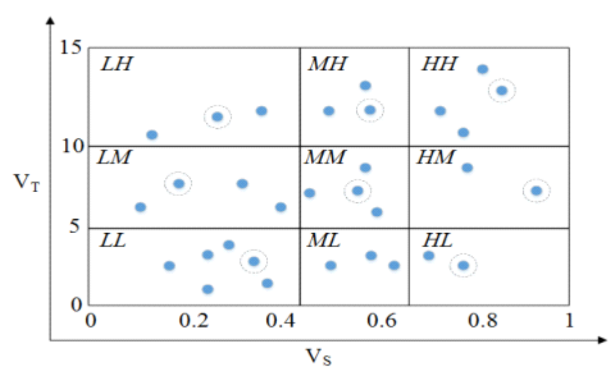

그림 2. "Collected-IR-Moving Camera Sequence" 범주의 후보 시퀀스에 대한 2차원 산포도. 점선 원은 선택된 테스트 시퀀스의 각 해당 등급의 대표로서 표시함
Fig. 2. 2-D Scatter plot of the candidate sequences from “Collected-IR-Moving Camera Sequence” category. The dashed circle indicates the selected test sequences as a representative of that class

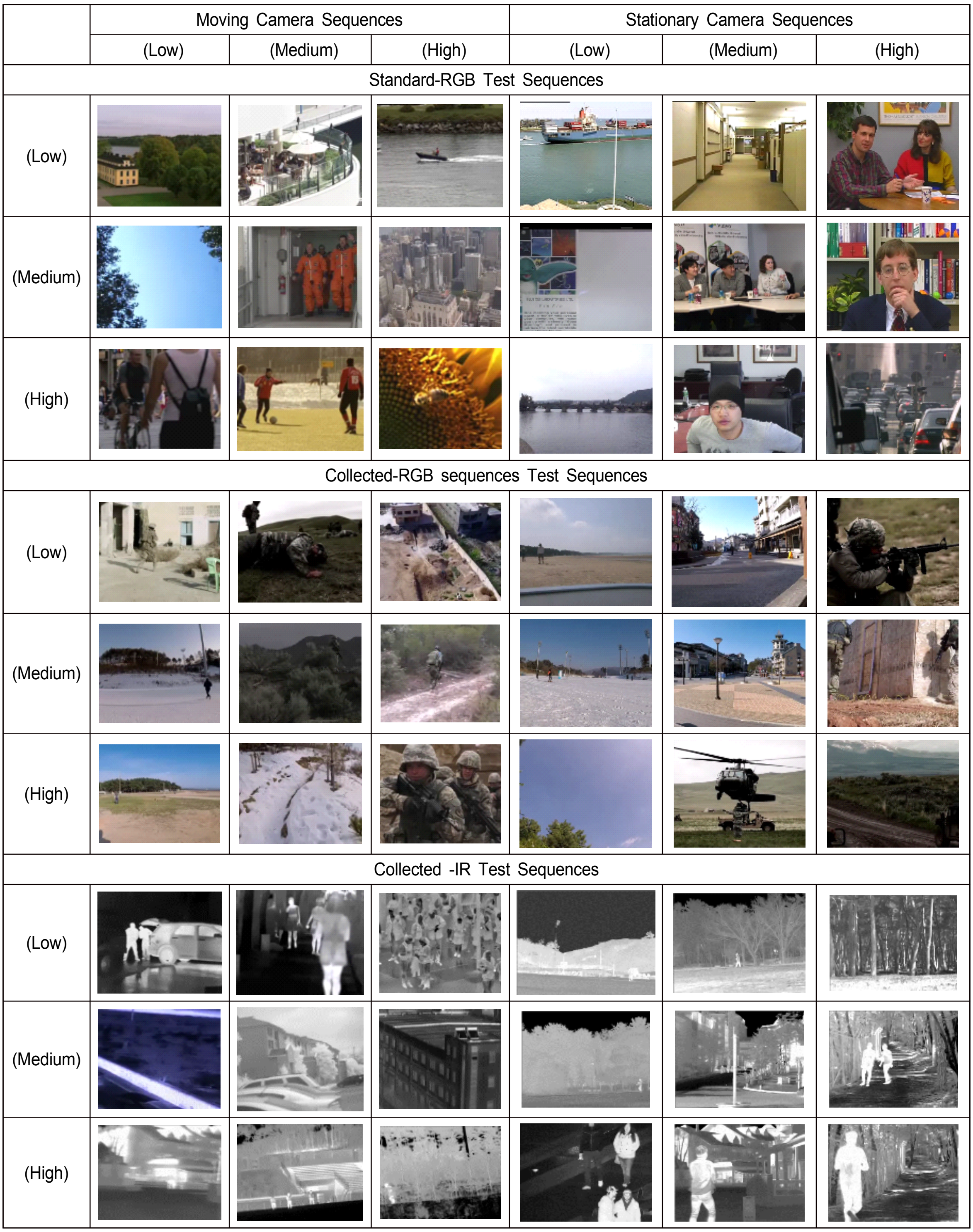

그림 3. 선택된 54개의 테스트 시퀀스
Fig. 3. Snapshots of the selected 54 test sequences

are selected for one category of video sequences, which ensure the variation in characteristics of the selected test sequences. This process is repeated for the rest of the five categories. Figure 2 shows a scatter plot of the candidate sequences taken from "Collected-IR-Moving Camera Sequence" category where the blue filled dots show the candidate sequence based on their spatial-temporal complexity value and the dashed circle indicates the selected representative sequences for the test set. Figure 3 shows the snapshots of the selected test sequences.

Ⅲ. Subjective Quality Evaluation

1. Subjective Quality Assessment

Subjective quality assessment aims to evaluate the perceived quality of videos under test by human observer, which is the best way for quality evaluation, because humans are the ultimate viewer of any video. According to ITU-R BT500-11 [9], at least 15 subjects are recommended for subjective quality assessment.

In this paper, there are 16 persons participated for the subjective quality assessment process and all of them have some basic ideas about image and video processing areas. The age range of the subjects are from 22 to 35 years old. Also for the environment setup to satisfy the laboratory conditions such as, laboratory luminance, viewing distance, viewing angle etc., we follow the recommendation of [9]. There are several subjective quality evaluation methodologies, defined by ITU-R BT500-11 [9] such as: Single-Stimulus (SS); Double-Stimulus Impairment Scale (DSIS); Paired Comparison (PC), etc. In this paper, we adopted the DSIS method for subjective quality assessment. Following DSIS, we present a reference sequence followed by an impaired sequence without repetition, where the impaired sequences are prepared by encoding with three target bit-rates i.e., 20, 45 and 65 Kbps against the reference test sequences. To avoid biased results accumulated by various factors, the presentation order of test sequences is pseudo-randomized in each session for different subjects. A presentation of the DSIS method is shown in figure 4. The reference sequence is presented during the T1 period of 10 seconds and to eliminate the after-image effect, a gray sequence is presented during the T2 period of 3 seconds. Then the corresponding impaired sequence is presented during the T3 period of 10 seconds and again to eliminate the after-image effect, a gray sequence is displayed during the T4 period of 7 seconds. Also, the subjects evaluate the perceived quality of each impaired sequences during the T4 period. It is to be noted that, before starting the evaluation session, a brief outline is given to educate the subjects. The discrete scale of the voting score for subjective video quality assessment of DSIS method has a range from 1 to 5 as suggested by [9]. In order to avoid eye stress, no session takes greater than 30 minutes. So the whole evaluation process is divided into three sessions for each subject and every subjects evaluate all of the impaired sequences. In a single session the actual number of video

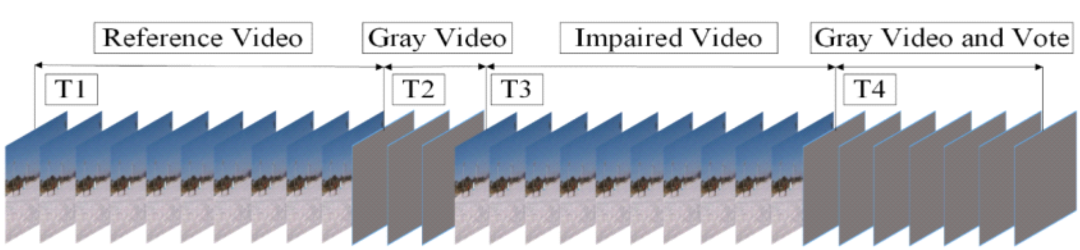

그림 4. 주관적 품질 평가를 위한 DSIS의 참조 및 손상된 시퀀스 제시
Fig. 4. Presentation of reference and impaired sequences in DSIS for subjective quality evaluation

sequences is 54 but 3 (control sequences) videos are repeated to check the biasness of the subjects, so total 57 video sequences are presented. Every session has 57 presentations and one presentation takes 30 seconds to complete. Therefore, it takes around 30 minutes to finish one session.

2. Data Processing and Analysis

After the subjective evaluation, we have our collected data that are used to calculate the MOS. However, there may be a few subjects who give an unnatural score to some videos compared to other subjects and can be defined as outliers. Outliers are not desired and should be removed before calculating the MOS. In this study, we use Z-Score which is very well known and widely used outlier removing strategy. Z-Score can be defined as:

$$Z = \frac{x - \overline{x}}{\sigma} \tag{1}$$

where $\overline{x}$ and σ are the mean and standard deviation of the data and can be defined as: $\overline{x} = \frac{1}{n}\Sigma_{i=1}^{n} x_i$ and $\sigma = \frac{\Sigma(x - \overline{x})}{n}$. More specifically, Z-Score measures how many standard deviations from the mean a score is. Z-Score ranges from -3 standard deviation up to +3 standard deviation and looks like a normal distribution. If any score is outside of this range, then it is considered as an outlier and removed from the data. Finally, we measure the MOS as follows:

$$MOS = \frac{1}{N}\Sigma_{i=1}^{N} S_i \tag{2}$$

where N is the number of subjects; and S_i is the i^{th} score

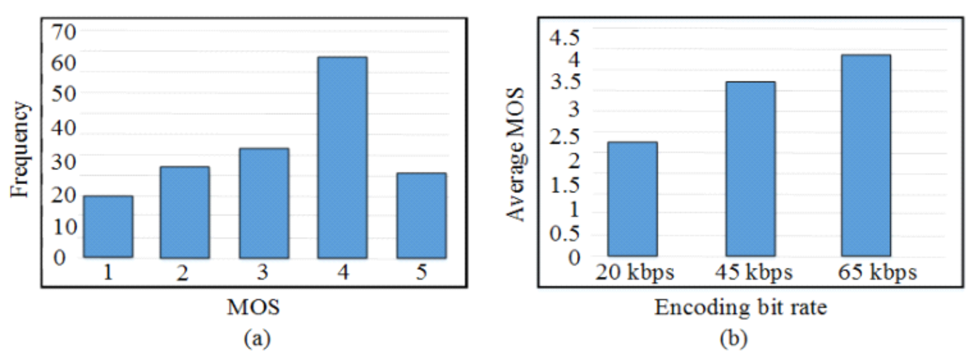

그림 5. 품질평가 MOS값의 주요 특성. (a) 수집된 점수의 히스토그램; (b) 각 비트 레이트에 대한 평균 MOS
Fig. 5. Primary characteristics of the found MOS values. (a) Histogram of the collected scores; (b) Average MOS for each bit-rate

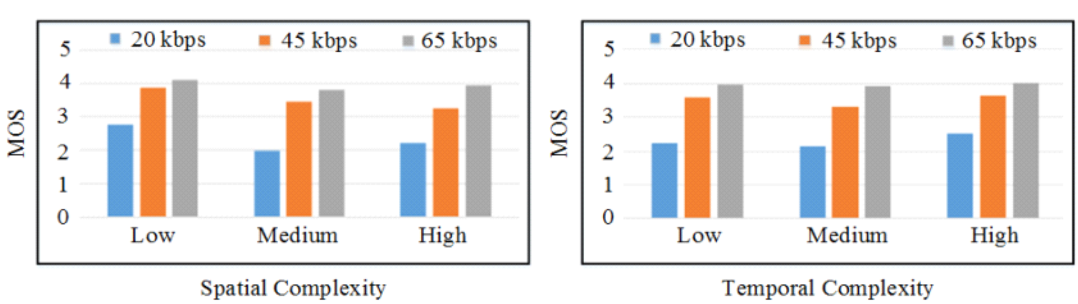

그림 6. 공간적, 시간적 복잡도에 기반한 MOS값
Fig. 6. MOS values based on spatial and temporal complexity

of an impaired sequence. After processing the raw data, we have the MOS values. Now we observe some primary characteristics of the collected MOS values. First, we measure the histogram of the scores as shown in Figure 5 (a), which indicates that the perceived quality of most of the sequences are good. Second, we measure the average MOS values for each of the target bit-rates that are used to encode the test sequences. The result is shown in Figure 5 (b), which indicates that the video sequences, even encoded with bit-rate of 45 kbps, still can satisfy the users in terms of perceived quality while utilizes a small amount of storage and a lower amount of bandwidth. Further, we analyze the MOS values based on spatial and temporal complexity. It is expected that the MOS value will be decreasing with the increase in spatial or temporal complexity. Figure 6 shows that the MOS is highest for low spatial and temporal complexity. But for high spatial and temporal complexity, the MOS is greater than the medium class, that is undesired. We found that there are few videos in high spatial and temporal classes that have very high MOS value, resulting in greater average MOS. Figure 7 shows the first frame of those videos where (a), (b) and (d) belongs to the high spatial class and (a), (c) and (e) belongs to the high temporal class. It can be seen that (a) consists of repeating textures, (e) consists of homeogenous regions and (d) has very low motion, which helps the encoder to achieve high prediction accuracy, resulting in a higher MOS. On the other hand, (b) and (c) contains a lot of high-frequency signals that are not very sensitive to HVS, thereby achieving high MOS.

Ⅳ. Comparison of Objective Quality Assessment Methods

In digital image processing, image quality assessment (IQA) is a fundamental part for instructing and optimizing the image and video applications, such as compression, transmission, super-resolution, restoration, etc. With the passing of time, many well established objective IQA met-

(a) (b) (c) (d) (e)

그림 7. 높은 공간적, 시간적 클래스에 크게 기여하는 첫 번째 프레임
Fig. 7. First frame of videos that are highly contributing in high spatial and temporal classes

표 2. 최신 이미지 및 비디오 품질평가 방법의 예측 정확도. 각 행에 대해 첫 번째, 두 번째 및 세 번째 성능을 가진 메트릭은 파란색, 빨간색 및 검은색 굵은 글씨로 강조 표시됨
Table 2. Prediction accuracy of the state-of-the-art image and video quality assessment methods. For each row, the first, second and third-ranked performances are highlighted by blue, red and black bold face texts

			FR-IQA			FR/RR-VQA			NR-VQA		
Metric	PSNR	SSIM	UQI	RFSIM	GMSD	VSI	SCQI	VSSCQI	SpEED	3DPSD	VIIDEO
SROC	**0.5730**	**0.6142**	**0.6028**	**0.5667**	**0.7446**	**0.6453**	**0.6191**	**0.6074**	**0.7669**	**0.6971**	**0.3806**
KROC	0.4182	0.4393	0.4369	0.4092	0.5535	0.4676	0.4574	0.4469	0.5702	0.5185	0.2602
PLCC	**0.6188**	**0.6581**	**0.6099**	**0.5959**	**0.7868**	**0.6821**	**0.6543**	**0.6452**	**0.8133**	**0.6978**	**0.3709**
RMSE	**0.7821**	**0.7495**	**0.7888**	**0.7994**	**0.6145**	**0.7279**	**0.7528**	**0.7601**	**0.5793**	**0.7130**	**0.9245**

rics have been developed which can automatically predict the quality of images. In our experiments, we have applied state-of-art full-reference IQA (FR-IQA) methods [10-16], full-reference video quality assessment (FR-VQA) methods [17-18] and no-reference video quality assessment (NR-VQA) method [19]. The IQA methods are applied for all of the frames and average is taken as the quality score. To measure the prediction accuracy, four well-accepted statistical indices Spearman Rank-Order Correlation coefficient (SROC), Kendal Rank-Order Correlation coefficients (KROC), Pearson Linear Correlation coefficients (PLCC) and Root Mean Square Error (RMSE) are used. All statistical indices have different meaning and they demonstrate the performance prediction from different aspects. The SROC, KROC measures the monotonicity between the estimated values by the objective quality assessment metrics and their subjective visual quality scores (MOS), PLCC measures the prediction accuracy and RMSE measures the prediction error. Usually, for quality assessment methods the SROC and KROC are considered as a representative performance measures. It is expected from a quality assessment metric is to achieve SROC, KROC and PLCC values close to 1 and RMSE values close to 0. In order to calculate the PLCC and RMSE, a logistic regression is adopted as a non-linear mapping function, that maps the estimated values of objective quality assessment methods to their measured subjective quality scores. The mathematical definition of the mapping function is:

$$f(x) = \beta_1 (\frac{1}{2} - (1 + exp^{\beta_2 (x - \beta_3) - 1}) + \beta_4 x + \beta_5 \qquad (3)$$

where x is the prediction score; $f(x)$ represents the corresponding subjective score and β_i denotes the parameters to be fitted. The prediction accuracy of objective image and video quality assessment methods are listed in Table 2 where the best three performances are highlighted by blue, red and black bold face texts. It is expected that in our proposed database, the VQA methods perform better than the IQA methods, since they utilize both the spatial and temporal information. From Table 2, it can be seen that the prediction accuracy for the majority of the methods are poor. Compared to other methods, the spatial efficient entropic differencing for image and video quality (SpEED) [17], that is a VQA method, outperforms other methods. Among IQA methods, the gradient magnitude similarity deviation (GMSD) [13], uses image gradients as features to predict the objective quality, shows competitive performance compared to SpEED and better than other two VQA methods 3DPSD [18] and VIIDEO [19]. 3DPSD, that estimates the quality of a distorted video by observing the changes in video dynamics in the frequency domain, also shows competitive results. However, the correlation is not satisfactory, hence we are in need of a suitable video quality assessment method for this kind of low-resolution surveillance videos.

Ⅴ. Conclusion

This paper has examined the perceptual quality assessment problem of low-resolution surveillance videos to use in ultra-low band transmission systems. A large number of candidate videos have been collected from real-world scenarios, using stationary and moving cameras for RGB and IR color formats. Then 54 test sequences have been selected by considering the spatial and temporal complexities, which largely affects the quality assessment process. H.265/HEVC video encoding technique has been applied with very low target bit-rates to get the impaired video sequences, on which the subjective quality scores have been collected via the DSIS method. The MOS values indicate that the videos even encoded with bit-rate of 45 kbps, can satisfy the human in terms of perceived quality. Moreover, we compared the state-of-the-art objective I/V-QA methods on the proposed database. Although, the SpEED, GMSD and 3DPSD methods achieve competitive correlation with the subjective MOS, still a better and dedicated quality assessment method is required to measure the objective quality of low-resolution surveillance video sequences.

참 고 문 헌 (References)

[1] G. J. Sullivan and J.-R. Ohm, “Recent developments in standardization of high efficiency video coding (HEVC),” *Proceeding of SPIE*, vol. 7798. Aug. 2010, paper 7798-30, https://doi.org/10.1117/12.863486.

[2] The Moving Picture Experts Group, https://mpeg.chiariglione.org/

[3] Test Materials to Be Used in Subjective Assessment of picture quality, *ITU-R BT.1210, International Telecommunications Union*, Feb. 2004.

[4] C. S. Won, D. K. Park, and S.-J. Park, “Efficient use of MPEG-7 edge histogram descriptor,” *Electronics and Telecommunications Research Institute Journal*, vol. 24, no. 1, pp. 23–30, Feb. 2002, https://doi.org/10.4218/etrij.02.0102.0103.

[5] K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack, "Study of Subjective and Objective Quality Assessment of Video", *IEEE Transactions on Image Processing*, vol.19, no.6, pp.1427-1441, June 2010, https://doi.org/10.1109/TIP.2010.2042111.

[6] P. V. Vu and D. M. Chandler, "ViS3: An Algorithm for Video Quality Assessment via Analysis of Spatial and Spatiotemporal Slices," *Journal of Electronic Imaging,* 23 (1), 01316, 2014, https://doi.org/10.1117/1.JEI.23.1.013016.

[7] C. G. Bampis, Z.Li, I. Katsavounidis, TY Huang, C. Ekanadham and A. C. Bovik, “Towards Perceptually Optimized End-to-end Adaptive Video Streaming,” *submitted to IEEE Transactions on Image Processing*, 2018, https://arxiv.org/abs/1808.03898.

[8] Gary J Sullivan, Jensrainer Ohm,Woojin Han, and ThomasWiegand, “Overview of the high efficiency video coding (HEVC) standard,” *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 22, no. 12, pp. 1649–1668, 2012, https://doi.org/10.1109/TCSVT.2012.2221191.

[9] Methodology for the Subjective Assessment of the Quality of TV Pictures, ITU-R BT.500-11, *International Telecommunications Union*, Jun. 2002.

[10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, Apr. 2004, https://doi.org/10.1109/TIP.2003.819861.

[11] Zhou Wang and A. C. Bovik, "A universal image quality index," in *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81-84, March 2002, https://doi.org/10.1109/97.995823.

[12] L. Zhang, L. Zhang and X. Mou, "RFSIM: A feature based image quality assessment metric using Riesz transforms," *2010 IEEE International Conference on Image Processing*, Hong Kong, 2010, pp. 321-324, https://doi.org/10.1109/ICIP.2010.5649275.

[13] W. Xue, L. Zhang, X. Mou and A. C. Bovik, "Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index," in *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684-695, Feb. 2014, https://doi.org/10.1109/TIP.2013.2293423.

[14] L. Zhang, Y. Shen and H. Li, "VSI: A Visual Saliency-Induced Index for Perceptual Image Quality Assessment," in *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4270-4281, Oct. 2014, https://doi.org/10.1109/TIP.2014.2346028.

[15] S. Bae and M. Kim, "A Novel Image Quality Assessment With Globally and Locally Consilient Visual Quality Perception," in *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2392-2406, May 2016, https://doi.org/10.1109/TIP.2016.2545863.

[16] A. F. M. S. Uddin, T. C. Chung and S. Bae, "Visual saliency based structural contrast quality index," in *IET Electronics Letters*, vol. 55, no. 4, pp. 194-196, Feb. 2019, https://doi.org/10.1049/el.2018.6435.

[17] C. G. Bampis, P. Gupta, R. Soundararajan and A. C. Bovik, "SpEED-QA: Spatial Efficient Entropic Differencing for Image and Video Quality," in *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1333-1337, Sept. 2017, https://doi.org/10.1109/LSP.2017.2726542.

[18] M. A. Aabed, G. Kwon and G. AlRegib, "Power of tempospatially unified spectral density for perceptual video quality assessment," *2017 IEEE International Conference on Multimedia and Expo (ICME)*, Hong Kong, 2017, pp. 1476-1481, https://doi.org/10.1109/ICME.2017.8019333.

[19] A. Mittal, M. A. Saad and A. C. Bovik, "A Completely Blind Video Integrity Oracle," in *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 289-300, Jan. 2016, https://doi.org/10.1109/TIP.2015.2502725.

저 자 소 개

A. F. M. Shahab Uddin

- 2015년 : 방글라데시 Islamic University, Information and Communication Engineering 학사
- 2016년 : 방글라데시 Islamic University, Information and Communication Engineering 석사
- 2017년 9월 ~ 현재 : 경희대학교 박사과정
- ORCID : https://orcid.org/0000-0003-1074-0515
- 주관심분야 : 이미지 품질 측정, 지각적 이미지 처리, 이미지 처리에서의 딥 러닝 및 역 문제

Mst. Sirazam Monira

- 2015년 : 방글라데시 Islamic University, Information and Communication Engineering 학사
- 2016년 : 방글라데시 Islamic University, Information and Communication Engineering 석사
- 2017년 9월 ~ 현재 : 경희대학교 석사과정
- ORCID : https://orcid.org/0000-0001-6932-5557
- 주관심분야 : 이미지 품질 측정, 딥 러닝 기반 이미지 처리, 데이터 확장

정 태 충

- 1980년 : 서울대학교 전자공학과 학사
- 1982년 : KAIST 컴퓨터과학 석사
- 1987년 : KAIST 컴퓨터과학 박사
- 1988년 ~ 현재 : 경희대학교 컴퓨터공학과 교수
- ORCID : https://orcid.org/0000-0001-7387-5113
- 주관심분야 : 기계학습, 메타서치, 로보틱스

김 동 현

- 2009년 2월 : 연세대학교 전기전자공학부 학사
- 2011년 2월 : 연세대학교 전기전자공학부 석사
- 2011년 6월 ~ 현재 : 국방과학연구소 제2기술연구본부
- ORCID : https://orcid.org/0000-0002-2136-5944
- 주관심분야 : 전술통신, 영상전송시스템, 데이터링크

최 증 원

- 1989년 2월 : 충남대학교 계산통계학과 학사
- 1993년 8월 : 충남대학교 계산통계학과(전산학) 석사
- 1997년 8월 : 충남대학교 전산학과 박사
- 1997년 7월 ~ 현재 : 국방과학연구소 수석연구원
- 2013년 9월 ~ 현재 : 과학기술연합대학원대학교 부교수
- ORCID : https://orcid.org/0000-0002-3642-2323
- 주관심분야 : 전술통신, 위성통신, 인지무선통신, 바이오통신, 정보융합 등

저 자 소 개

전 기 남

- 2005년 2월 : 한양대학교 전자컴퓨터공학부 학사
- 2007년 2월 : 한양대학교 컴퓨터공학과 석사
- 2007년 1월 ~ 현재 : LIG 넥스원 수석연구원
- ORCID : https://orcid.org/0000-0003-1414-5297
- 주관심분야 : 센서 네트워크, 신호처리, 정보융합, 영상처리, 영상압축

배 성 호

- 2011년 : 경희대학교 전자공학과 학사
- 2012년 : KAIST 전기및전자공학 석사
- 2016년 : KAIST 전기및전자공학 박사
- 2016년 ~ 2017년: MIT 컴퓨터과학 박사후과정
- 2017년 ~ 현재 : 경희대학교 컴퓨터공학과 조교수
- ORCID : https://orcid.org/0000-0003-2677-3186
- 주관심분야 : 심층신경망 모델 압축/해석/탐색, 비디오코딩, 이미지 처리에서의 역문제