

3. PSNR CURVES

Selecting a specific spatial resolution and bitrate combination to transmit video on MobileASL is important because there are tradeoffs with computational complexity, video quality, and resource availability on a cell phone such as battery life and data rate consumption. Larger video resolutions and higher bitrates result in higher video quality at the expense of increased computational power to transmit the data in real time. Before we can investigate how resource allocation is affected by video transmission, we need to determine at which bitrates and spatial resolutions we can get high enough video quality for intelligible conversations.

Despite the fact that PSNR may not be suitable for measuring subjective video quality, it still is a reasonable measurement of video quality when used across the same content [25]. We calculated PSNR of two different spatial resolutions (192×140 and 320×240 pixels) and 15 bitrates (10-150 kbps in increments of 10 kbps) of the same ASL video. The smaller spatial resolution was transmitted at 192×140 pixels and then enlarged and displayed at 320×240 pixels using bilinear interpolation [27] before PSNR was calculated.

The same 12-second video clip of a local deaf woman signing at her natural signing pace with a stationary background was used in the calculation of PSNR. The video was recorded at 320×240 pixels at 15 fps. Duplicate videos were created at the smaller spatial resolution before calculating the PSNR. The x264 codec, an open source version of H.264 codec, was used to compress the videos at each spatial resolution and bitrate combination [3,24]. As Figure 2 demonstrates, the PSNR values for each spatial resolution increase monotonically with increasing bitrate.

We found that the PSNR curves demonstrated a crossover point where, at lower bitrates (40 kbps and below), the smaller spatial resolutions had higher PSNR values than the larger spatial resolution. Visual inspection of the same ASL video (displayed at the same size) transmitted at lower bitrates (10-40 kbps) showed more blocky artifacts in videos sent at 320×240 pixels than at 192×144. The crossover in the PSNR plots occurred because at very low bitrates, the higher resolution video is quantized more heavily and thus has very poor visual quality (such as blockiness and loss of fine details). The same videos at lower spatial resolutions are not quantized as heavily which results in higher measured video quality. As bitrates increase, the higher resolution has higher measured video quality than the smaller spatial resolutions. This is due to blurriness from enlarging the video. The crossover of PSNR curves has been found in other video compression techniques [16,19,29], but the results, to our knowledge, have not been used to evaluate human comprehension, which, along with subjective quality measures, is the focus of our online survey.

4. ONLINE SURVEY METHOD

From a technological perspective, transmitting video at the smaller spatial resolution and at the lowest bitrates takes the least amount of computational power and resources; however, without feedback from users, we cannot confirm that sign language communication with this video is intelligible.

We created and deployed a national three-part online survey to investigate user preferences and comprehension when varying the bitrates (10-60 kbps in increments of 10 kbps) and spatial resolutions (192×144 and 320×240) of ASL video. We did not consider bitrates higher than 60 kbps since the larger spatial

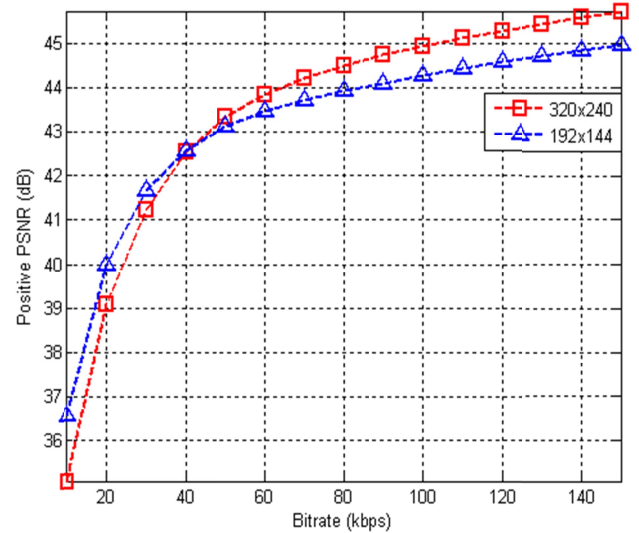


Figure 2: PSNR(dB) vs. Bitrate (kbps) for spatial resolutions displayed at 320×240 pixels. Higher PSNR means higher objective video quality. Whether it means higher subjective perception of quality is a topic of this research.

resolution always had higher video quality than the smaller spatial resolution.

The online survey began by asking participants to self-report their fluency in ASL. The survey asked different questions depending on the response to this question. Part 1 was a paired-comparison experiment which investigated the subjective video quality preferences of ASL speakers and non-ASL speakers (see Figure 3). Part 2 was a single-stimulus experiment which examined comprehension of ASL video of varying bitrates and spatial resolutions (ASL speakers only) (see Figures 4 and 5). Finally, part 3 asked demographic questions.

To determine how subjective video quality preference differs between ASL speakers and non-ASL speakers, it was important to get an equal number of ASL and non-ASL speaking respondents. We selected an online survey over a laboratory study because an online survey is accessible to most people with Internet access, so more respondents could be included from across the nation.

4.1 Videos Used in Online Survey

4.1.1 Videos in Part 1

The same 12-second video clips used to measure PSNR (see section 3, above) of ASL video were used in part 1 of the survey. A 12-second video duration was used because it was long enough for respondents to make a video preference selection while keeping the overall survey manageable to complete in 4-7 minutes. Recall that all videos were transmitted at their respective spatial resolution (192×144 and 320×240) at varied bitrates, and then displayed at 320×240 pixels (with the smaller spatial resolution enlarged using bilinear interpolation).

4.1.2 Videos in Part 2

Twelve different video clips of the same local deaf woman signing different short stories at her natural signing pace were used.

- Conference on Computers and Accessibility*, (pp. 163-170). Tempe, AZ.
- [7] Ciaramello, F. & Hemami, S. (2011, January). Quality versus Intelligibility: Studying Human Preferences for American Sign Language Video. *Proceedings in SPIE Volume 7865, Human Vision and Electronic Imaging*.
- [8] Feghali, R., Speranza, F., Wang, D., & Vincent, A. (2007, March). Video Quality Metric for Bit Rate Control via Joint Adjustment of Quantization and Frame Rate. *53*(IEEE Transactions on Broadcasting).
- [9] Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association* *32* (200), 675-701
- [10] Fring. Retrieved 4 29, 2011, from <http://www.fring.com/>
- [11] Girod, B. (1993). What's wrong with mean-squared error? *Digital images and human vision*, 207-220.
- [12] Hooper, S., Miller, C., Rose, S., & Veletsianos, G. (2007). The Effects of Digital Video Quality on Learner Comprehension in an American Sign Language Assessment Environment. *Sign Language Studies*, *8*(Sign Language Studies), 42-58.
- [13] Huynh-Thu, Q., & Ghanbari, M. (2008). Scope of validity of PSNR in image/video quality assessment. The Institution of Engineering Technology.
- [14] ITU. (September 1999). p.910: Subjective video quality assessment methods for multimedia applications.
- [15] Jumiski-Pyykko. (2005). Evaluation of Subjective Video Quality of Mobile Devices. *Multimedia Proceedings of the 13th annual ACM international conference on Multimedia*. Singapore.
- [16] Lin, W., & Dong, L. (2006, September). Adaptive Downsampling to Improve Image Compression at Low Bit Rates. *IEEE Transactions on Image Processing*. *15*.
- [17] Masry, M., & Hemami, S. (2003, January). CVQE: A metric for continuous video quality evaluation at low rates. *Proceedings in SPIE: Human Vision and Electronic Imaging*.
- [18] Nemethova, A., Ries, M., Zavodsky, M., & Rupp, M. (2006). PSNR-Based Estimation of Subjective Time-Variant Video Quality for Mobiles. *Measurement of Audio and Video Quality in Networks*.
- [19] Nguyen, V., Tan, Y., Z., & Lin, W. (2006). Adaptive Downsampling/Upsampling for better video compression at low bit rate. *IEEE International Symposium on Circuits and Systems*.
- [20] *Purple VRS on Your Devices*. (Purple Communications.) Retrieved 7 31, 2011, from <http://www.purple.us/>
- [21] *Qik*. (Qik, Inc.) Retrieved 4 29, 2011, from <http://qik.com/>
- [22] Reiter, U., & Korhonen, J. (2009). Comparing Apples and Oranges: Subjective Quality Assessment of Streamed Video with Different Types of Distortion. (IEEE).
- [23] Radaelli-Sanchez, Baraniuk, R. (2010). Gibbs's Phenomena. <http://cns.org/content/m10092/latest>
- [24] Richardson, I. (2004). vocdex: H.264 tutorial white papers.
- [25] Thu, H., & Ghanbari, M. (2008). Scope of Validity of PSNR in image/video quality assessment. *44*.
- [26] Tran, J. J., Johnson, T. W., Kim, J., Rodriguez, R., Yin, S., Riskin, E., Ladner, R., Wobbrock, J., (2010). A Web-Based User Survey for Evaluating Power Saving Strategies for Deaf Users of MobileASL. *Proceedings of ASSETS 2010: The 12th International ACM SIGACCESS Conference on Computers and Accessibility*. Orlando, FL.
- [27] Vision Systems Design. Retrieved 5 03, 2011. Understanding image-interpolation techniques.
- [28] VQEG. (2003, September). Final report from the video quality experts group on the validation of objective models of video.
- [29] Wang, R. Chien, M., Chang, P. Adaptive Down-Sampling Video Coding. VQEG. (2010). *Proceedings in SPIE 7542*.
- [30] Wang, Z., Bovik, A., & Lu, L. (2002). Why is Image Quality Assessment so Difficult? *IEEE Acoustic, Speech, and Signal Processing*.
- [31] Wang, Z., Lu, L., & Bovik, A. (2004, February). Video quality assessment based on structural distortion measurement. *19*(Signal Processing: Image Communication special issue on Objective video quality metrics), 121-132.
- [32] Wiegand, T., Schwarz, H., Joch, A., Kossentini, F., & Sullivan, G. (2003). Rate-constrained coder control and comparison of video coding standards. *IEEE Transactions Circuits Systems Video Technology*, *13*(7), 688-703.
- [33] Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin* *1* (6), 80-83.
- [34] Winkler, S., & Faller, C. (2005). Audiovisual Quality Evaluation of Low-Bitrate Video. *SPIE/IS&T Human Vision and Electronic Imaging*, *5666*, pp. 139-148. San Jose.
- [35] Winkler, S., & Faller, C. (2005). Maximizing Audiovisual Quality at Low Bitrates. *Workshop on Video Processing and Quality Metrics for Consumer Electronics*. Scottsdale, AZ.
- [36] *ZVRS*. (ZVRS Communication Service for the Deaf, Inc.) Retrieved 7 31, 2011, from <http://www.zvrs.com/z-series/>