A Web-Based User Survey for Evaluating Power Saving Strategies for Deaf Users of MobileASL

Jessica J. Tran[1], Tressa W. Johnson[2], Joy Kim[3], Rafael Rodriguez[3]
Sheri Yin[3], Eve A. Riskin[1], Richard E. Ladner[3], Jacob O. Wobbrock[2]

[1]Electrical Engineering
University of Washington
Seattle, WA 98195 USA
{jjtran, riskin}@ee.washington.edu

[2]The Information School
DUB Group
University of Washington
Seattle, WA 98195 USA
{tressaj, wobbrock}@uw.edu

[3]Computer Science & Engineering
University of Washington
Seattle, WA 98195 USA
{ladner}@cs.washington.edu
{jojo0808, rodriraf, shericn}@uw.edu

ABSTRACT
MobileASL is a video compression project for two-way, real-time video communication on cell phones, allowing Deaf people to communicate in the language most accessible to them, American Sign Language. Unfortunately, running MobileASL quickly depletes a full battery charge in a few hours. Previous work on MobileASL investigated a method called *variable frame rate (VFR)* to increase the battery duration. We expand on this previous work by creating two new power saving algorithms, *variable spatial resolution (VSR),* and the application of both VFR and VSR. These algorithms extend the battery life by altering the temporal and/or spatial resolutions of video transmitted on MobileASL. We found that implementing only VFR extended the battery life from 284 minutes to 307 minutes; implementing only VSR extended the battery life to 306 minutes, and implementing both VFR and VSR extended the battery life to 315 minutes. We evaluated all three algorithms by creating a linguistically accessible online survey to investigate Deaf people's perceptions of video quality when these algorithms were applied. In our survey results, we found that VFR produces perceived video choppiness and VSR produces perceived video blurriness; however, a surprising finding was that when both VFR and VSR are used together, they largely ameliorate the choppiness and blurriness perceived, *i.e.*, they each improve the use of the other. This is a useful finding because using VFR and VSR together saves the most battery life.

Categories and Subject Descriptors
K.4.2. [**Social Issues**]: Assistive technologies for persons with disabilities; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – Video.

General Terms
Performance, Experimentation, Human Factors.

Keywords
Battery power consumption, encoding algorithms, video compression, web-based user survey, mobile phones, American Sign Language, Deaf Community, Deaf Culture.

	Question 1: I notice sections of this video are choppy.	Question 2: The choppy sections of this video are distracting.	Question 3: I notice sections of this video are blurry.	Question 4: The blurry sections of this video are distracting.
Strongly Agree	○	○	○	○
Somewhat Agree	○	○	○	○
Neutral	○	○	○	○
Somewhat Disagree	○	○	○	○
Strongly Disagree	○	○	○	○
Not Applicable		○		○

Figure 1: Screen shot of one video from our online survey for Deaf people.

1. INTRODUCTION
Current cell phone technology in the United States does not support real-time video conversations, and thus does not allow Deaf[1] people to communicate in the language most accessible to them, American Sign Language (ASL)[2]. For years, Deaf people have utilized the text messaging capabilities of cell phones; however, since text is a representation of an auditory language to which they do not have access, this means they are effectively

[1] Using capital "Deaf" is accepted practice when referring to members of Deaf Culture, while lower case "deaf" is used when referring to an individual with hearing loss.

[2] American Sign Language (ASL) is the signed language indigenous to the United States and English-speaking Canada. Research has shown the linguistic framework of ASL to be independent of English [20].

communicating in their second language. The MobileASL project [2] has developed a software video codec which allows very low bit rate coding (under 30 kilobits/second) to transfer real-time video on the current cellular network in the United States. This enables users to communicate in ASL while enjoying the advantages of a mobile phone. The HTC TyTN II cell phone was selected to run MobileASL because it has a front-facing camera, allowing users to sign and see with whom they are communicating. Unfortunately, running MobileASL has high power consumption rates that drastically deplete one full battery charge from 40 hours to 284 minutes [30]. A large contribution to the battery drain is the computational complexity of transmitting video at 10-12 frames per second (fps) in real time. Previous work on MobileASL investigated a method called *variable frame rate (VFR)* to increase the battery duration. We expand on this previous work by creating two new power saving algorithms, *variable spatial resolution (VSR),* and the application of both VFR and VSR. These algorithms extend a full battery charge by altering the temporal and/or spatial resolutions of video transmitted on MobileASL. These algorithms utilize activity recognition [8], which identifies signing and not-signing parts of a conversation. A conversation in ASL, as in English, involves times when a message is being conveyed (signing/talking), and attending to the message conveyed (not-signing/just listening). In our battery power study, we investigated the ideal case to save battery power, namely when the power saving algorithm is constantly applied, *i.e.*, a person is just listening to the conversation. Implementing only VFR extended the battery life from 284 minutes to 307 minutes; implementing only VSR extended the battery life to 306 minutes; and applying VFR and VSR together extended the battery life to 315 minutes.

We conducted a web-based user survey to investigate how Deaf users experience and feel about degradation of video quality in exchange for longer battery life on the phone. Comprehension of video content was not the focus of this survey. Figure 1 is a screen shot of one video from our online survey. We found that VFR produces perceived video choppiness and VSR produces perceived video blurriness; however, a surprising finding was that when VFR was on, the use of VSR significantly lowered the perceived choppiness of the video ($F(1,591)=8.09$, $p<.01$) and when VSR was on, the use of VFR significantly lowered the perceived blurriness of the video ($F(1,591)=18.99$, $p<.001$). These findings and others are presented in detail in our results.

The main contributions of this paper are presenting two new algorithms that extend a mobile video-enabled phone's battery life; presenting a method for creating an online survey intended for Deaf people; and discovering that using VFR and VSR together reduces the perceived choppiness and blurriness introduced by the individual algorithms.

2. RELATED WORK

Linguistic research has shown that ASL is not a visual code for English [20]; ASL has a distinct, unrelated grammar and lexicon that has developed over time in the context of a deaf community. There are conversational similarities between ASL and English, such as multiple people "holding the floor" at once [11] and feedback through back-channeling [12] (the latter refers to one person acknowledging understanding to the other, which could take the form of a muttered "uh-huh" in English or a head nod in ASL). Since MobileASL users may "sign over one another," we investigated if applying different power saving algorithms negatively impacted perceived video quality by modifying the temporal and/or spatial resolution of the video when one user is not-signing. Video quality perception that varies temporally and/or spatially depending on video content has not, to our knowledge, been studied.

Previous research has found that hand and face movements are key linguistic features of ASL that contribute to the intelligibility of a message [31]. Peripheral low-resolution vision is a key component in the perception of movement. Muir and Richardson [22] explored the eye movement patterns of Deaf people as they viewed sign language in video and then applied their findings to the design of video communication systems. Their findings concluded that a Deaf viewer's focus is placed on the facial region of a signer in order to pick up the small detailed movements in the signer's facial expression and lip shapes. This region is of interest because it conveys linguistic information to the receiver. Our work explores the human perception of video quality when different power saving algorithms are applied during not-signing sections of a conversation.

Another research topic of interest is the effect of frame rate on sign language comprehension. Johnson and Caird [24] discovered that 1 and 5 frames per second (fps) were enough for beginners to learn ASL from ten videos, each containing one sign. Sperling *et al.* [29] also found considerable reduction in comprehension from 10 to 5 fps, slight reduction in comprehension from 15 to 10 fps, and insignificant difference in comprehension from 30 to 15 fps. The videos used in our online survey contained "conversationally-paced" signing with many quickly produced signs and users who are experts in sign language. Previous research associated with MobileASL [8] found that transmitting 1 fps during not-signing sections of a conversation was sufficient for comprehension. Therefore, during our investigation of alternative power saving algorithms, we use 1 fps as the transmission rate for the manipulated temporal resolution of video. Our work creates a new power saving algorithm by combining the 1 fps transmission rate with an algorithm that alters the spatial resolution of video.

2.1 Surveys for Deaf Participants

Instructions in both English text and ASL videos have not, to our knowledge, been used in web-based user surveys intended for Deaf participants. Previous studies [4,17] have been conducted to examine electronic communication among the Deaf population, but the medium in which researchers primarily chose to gather data from Deaf participants was based on English text. Hogg *et al.* [17] researched the use of communication technology, gathering data from Deaf participants through an online English text based survey. In the analysis of their responses, they recognized the limitations of their survey due to a "high proportion" of participants that did not complete the survey. Hogg *et al.* suggested that the participants' variance in reading levels may have contributed to the incomplete surveys. They further suggested that an ASL version of their survey might have produced better results.

A study conducted by Akamatsu *et al.* [4] used a text-based survey to collect data for their investigation of two-way text messaging between Deaf high school students and their hearing parents. When the researchers reviewed the survey results, they recognized that the textual surveys were not linguistically accessible, and determined that interviews conducted in sign language were necessary to ensure complete and accurate responses from their Deaf participants. Our online survey is unique in that we attempted to create a linguistically accessible survey by presenting instructions in both English text and ASL videos.

2.2 Power Savings

Previous research conducted by Cherniavsky *et al.* [9] on the MobileASL project investigated our first power saving algorithm, *variable frame rate (VFR)*. Their research used an older version of MobileASL which sent larger QCIF (176×144 pixels) frames at a lower frame rate of 7-8 fps (this version was not used in our survey.) VFR manipulates the temporal resolution of the transmitted video to save computational resources. VFR uses activity recognition [7], which employs baseline differencing to calculate the sum of absolute differences of the luminance component of consecutive video frames:

$$d(k) = \sum_{i,j \in I(k)} | I_k(i,j) - I_{k-1}(i,j) |$$

In the equation above, (i,j) is the pixel coordinates within the k^{th} image frame I_k.

Frames that are classified as signing may contain a lot of activity such as fast movement in the hands or face, resulting in large pixel differences. Frames that are classified as not-signing have small pixel differences due to little change in background or movement in the hands and face by the user. If the difference between each frame is above a certain threshold, then frames are classified as signing, otherwise not-signing. When a not-signing frame is identified, VFR reduces the frame rate from 7-8 fps down to 1 fps, which produces a choppy video. Cherniavsky *et al.*'s [9] VFR implementation using the older version of MobileASL extended the battery life from 150 to 218 minutes. The longer extension of battery life is a result of VFR dropping frames that are 64% larger than the current MobileASL implementation (frame size 96×96 sent at 10-12 fps). Figure 2 demonstrates how video frames are reduced during the not-signing portions of a conversation.

Cherniavsky *et al.* [9] evaluated VFR in a laboratory setting. Fifteen participants fluent in ASL were recruited. The goal was to measure comprehensibility of conversations while the VFR algorithm was applied during the not-signing sections of a conversation. The objective measurements of this study included number of requests for repetition (repair requests) [29], number of turns associated with repair requests, number of conversational breakdowns, and the speed of finger spelling. Their findings revealed that when VFR was on, participants felt they had to guess at what was being said more frequently than when VFR was off. Applying the VFR algorithm also resulted in more repair requests, took more turns to correct the request, and resulted in more conversational breakdowns. These results prompted us to find alternative power saving algorithms to extend the battery life.

The objective of our online survey was to determine whether participants perceived changes in video quality during the application of our power saving algorithms, and if so, whether they find the change in video quality distracting. Our web-based survey is different from previous work because it was not conducted in a laboratory which allowed for greater recruitment of participants. Our goal was to investigate video perception rather than video intelligibility. Objective measures such as repeated request and conversational breakdowns were not investigated because participants viewed pre-recorded videos of one sided conversations (see section 5.3 for survey content.)

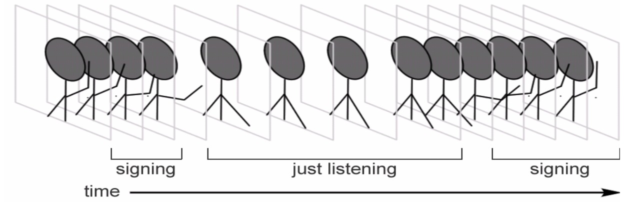

Figure 2: Depiction of variable frame rate algorithm. The frame rate decreases when the signer is not-signing, resulting in "choppy" video quality.

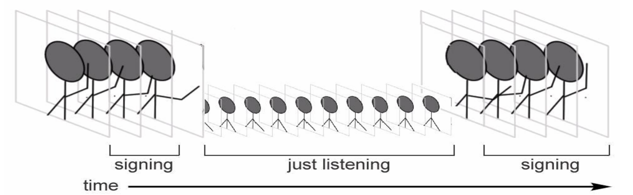

Figure 3: Depiction of variable spatial resolution algorithm. The not-signing frames downsample to 1/4 the original frame size, resulting in "blurry" video quality. (The image size shown on the mobile device is kept the same after downsampling.)

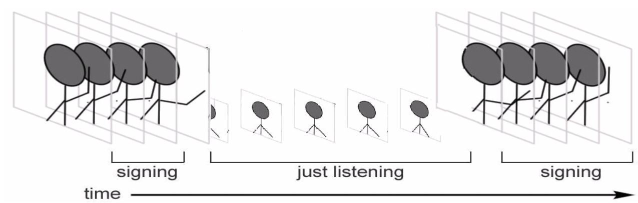

Figure 4: The implementation of VFR and VSR during not-signing portions of a conversation. The resulting output is a combination of blurry and choppy video.

3. NEW POWER SAVING ALGORITHMS

We present two new power saving algorithms that utilize activity recognition [9] to extend battery resources when using MobileASL (frame size 96×96 pixels sent at 10-12 fps).

3.1 Variable Spatial Resolution (VSR)

The first new power saving algorithm is called *variable spatial resolution (VSR)*. It is similar to VFR in that it uses the identification of signing or not-signing frames, but in this case, it *downsamples* the captured video frames. When the person is not-signing, the frame rate is held constant at 10-12 fps, but the frame size is downsampled to 1/4 of the original size before being sent to the video encoder. After the frame is transmitted, the receiving phone's decoder enlarges the downsampled frame to its original size, so the video appears blurry. Figure 3 demonstrates the implementation of VSR for not-signing or "just listening" portions of a conversation.

3.2 Combination of VFR and VSR

The second new power saving algorithm was the combination of both VFR and VSR. Intuitively, combining the two methods should produce further savings. Sending less data per frame along with fewer frames per second is computationally less intensive than applying either algorithm alone. When the user is not-signing, the frame rate is reduced from 10-12 fps to 1 fps and the frame size is reduced to 1/4 of the original size. This produces both a choppy and blurry video. Figure 4 depicts the combination of VFR and VSR.

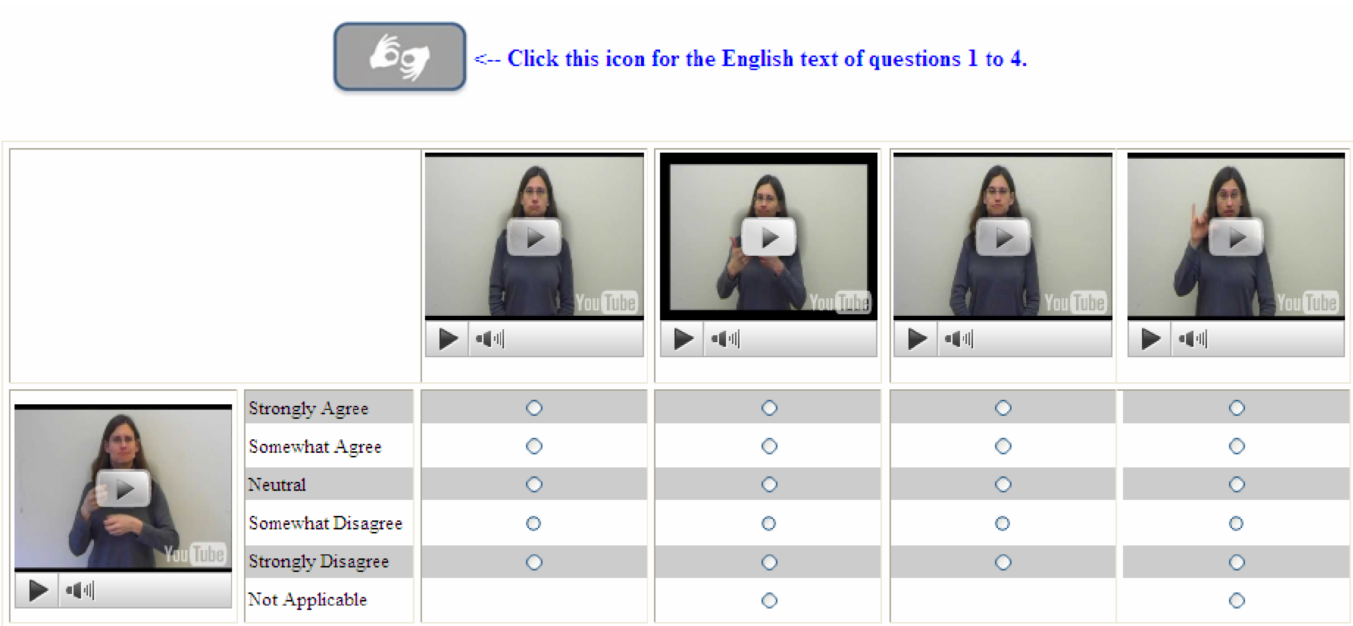

Figure 5: Screen shot of ASL video interpretation of survey questions and 5-point Likert scale.

3.3 Battery Power Results

Running MobileASL without any power saving algorithms consumes 76% of the phone's CPU and a full battery charge lasts on average 284 minutes. Implementing only VFR extended the battery life to 307 minutes and lowered the CPU usage to 26%. Implementing only VSR extended the battery life to 306 minutes and lowered the CPU usage to 32%. Applying the VFR and VSR algorithms together extended the battery life to 315 minutes and lowered the CPU usage to 10%.

4. METHODOLOGY FOR CREATING AN ONLINE SURVEY FOR THE DEAF

We created a web-based user survey with video prompts to evaluate the human perception of video quality when the three different power saving algorithms were applied. As we noted in part of our review of related work (section 2.1), it is nontrivial to create survey instruments for Deaf people, especially when English text must be relied upon. In this section, we share our lessons learned from creating an accessible online survey with video prompts for Deaf people.

There are two opposing conceptualizations of deafness, each with a unique impact on the design of a survey and the way in which it is received by Deaf participants. The first defines deafness as a pathological condition, while the second views deafness as a social identifier. The pathological model focuses on people's audiological status and considers deafness a medical condition requiring treatment. This perspective classifies people with hearing loss as "disabled" or "handicapped," and is marked by negative stereotypes and prejudice [14,23]. Under this paradigm, deafness is perceived as the dominant quality of a group of people who share a "condition."

The social model, in contrast, holds that Deaf people are disabled more by their interactions with hearing people than by the physical condition that determines their perception of sounds. This view recognizes the linguistic [20,21] and sociological [25,26] research that has identified ASL as a unique language distinct from English, and Deaf Culture as a legitimate culture distinct from the mainstream.

Given the historical dominance of the pathological view of deafness [19], designing an online survey that demonstrated respect for the language and culture of Deaf people was deemed of paramount importance. Taking into consideration both the values identified as defining characteristics of Deaf Culture, and the recorded experiences of deaf individuals who do not identify themselves as members of that culture, we identified two issues requiring explicit attention: linguistic accessibility, and respect for the autonomy and intelligence of the Deaf individual.

4.1 Online Survey Questions and Layout

To encourage high survey participation, we carefully considered the formulation of the questions asked and the layout of these questions. We used a 5-point Likert scale to gather participants' responses. We consulted the marketing director of Sorenson Communications, which conducts surveys of its deaf customers. He provided two suggestions for improving the survey's linguistic accessibility. The first suggestion was to simplify the survey questions for only "yes" or "no" responses. The disadvantage of this method was that we would have difficulty drawing in-depth conclusions from the results, so we opted out of this advice. The second suggestion was to tailor the questions so that they could be answered using specific responses such as:

- The video is very choppy.
- The video is moderately choppy.
- The video is slightly choppy.
- The video is not choppy at all.

The disadvantage of this method was that it made our survey English text heavy, therefore we also opted out of this advice. However, while investigating alternative ways to present survey questions we found that Allen *et al.* [5] suspect that a survey utilizing a horizontal layout may require additional interpretation by Deaf participants. They theorize that English speakers may intuitively recognize the spatial relationship between points on a horizontal Likert scale; in contrast, ASL, a nonlinear language, does not utilize this model of spatial relationships; therefore ASL speakers may not intuitively assign a horizontal layout the same implicit meaning as would English speakers. In response these researchers suggested that vertical Likert scales may be more accessible to Deaf participants, with the scalar zero point at the bottom and the end point at the top, mirroring spatial relationships found in the environment [5]. Therefore we adopted a vertical Likert scale layout as Figure 5 demonstrates.

4.2 American Sign Language Instructional Videos

Ensuring the accessibility of an online survey is paramount to its success. Three factors were taken into consideration with regard to the accessibility of our survey: the intended audience of Deaf signers; linguistic research determining the grammar and lexicon of ASL distinct from that of English [20,25]; and the value Deaf Culture places on both linguistic accessibility and self-determination [20]. We decided to provide an alternative to textual English by incorporating *ASL instructional videos*, as shown in Figure 5, to both increase accessibility and demonstrate the researchers' respect both for the individual participants and for Deaf Culture. Presenting the survey in two languages widened our audience to include both ASL signers and those who prefer to communicate visually (potential MobileASL users) but who are not fluent in ASL (example: late-deafened individuals.)

Neither words nor signs have absolute equivalents in other languages; what makes ASL/English interpretation possible is that both languages have the capacity to express identical meanings. The process of interpreting our survey in ASL began with analyzing the text for explicit and implicit meaning, English-based discourse patterns, and cultural influences. We then composed an interpretation that was equivalent in meaning while utilizing ASL-based discourse patterns and cultural influences. This process of interpreting the survey's English text into ASL was undertaken by the second author, a hearing, nationally certified[3] interpreter with over 10 years of professional experience.

5. ONLINE SURVEY OVERVIEW

Our online survey determined if users could detect the changes in video quality caused by each algorithm, especially during the not-signing portions of a conversation. A web-based survey was selected over a laboratory study because of our desire for greater numbers of participants who may be distant from our research site. To gather a large participation response, we partnered with Sorenson to send a press release of our survey to a random selection of 8,000 of their subscribers across the nation. To further promote our survey, we created a MobileASL Facebook group (which currently has over 232 members) and advertised this survey to a number of Facebook groups with members who are Deaf or interested in ASL. We also placed an ad in *Deafdigest* [1] to promote this online survey.

5.1 Survey Design

The design of the survey was a 2×2 within-subjects factorial design. The two factors were our two encoding schemes (VFR, VSR), each with two levels, "on" or "off." Figure 6 depicts the combinations of each factor and its levels.

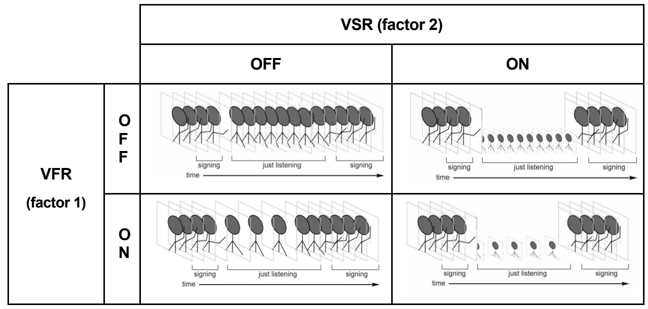

Figure 6: Combinations of factors and levels within our online survey.

5.2 Videos Used in Online Survey

Videos shown on a computer screen may appear differently than they would on a mobile phone due to differences in screen resolution, color mapping, and the decoder used by the media player. These issues were accounted for when creating the videos presented in our survey. To accurately represent mobile phone video on the computer screen, the videos were recorded with the video camera on the cell phone. We recorded three individual conversations of two local deaf women and a man, each signing at their own natural signing pace. The content of their conversations includes asking everyday questions such as how they are doing and what they did on the weekend. The recorded conversations were then encoded using H.264, which is the same encoder used by MobileASL. Afterwards, the encoded video was converted to MPEG-4 using a publicly available converter [18] that does not add additional artifacts. It was essential that additional artifacts not be added to the encoded video because it would interfere with perceived quality of the video. The Apple QuickTime media player [3] was used to play the videos on the computer screen.

5.3 Survey Content

Each participant was randomly assigned to view one of three videos of a person signing in ASL. The content of the video was a one-sided conversation with an equal amount of signing and not-signing. The assigned video was shown four consecutive times, but each time a different power savings algorithm was applied (VFR, VSR, VFR and VSR, or none). The participants did not know which encoding algorithm was applied; they were only told that there may be changes to the video quality, but not when, where, how, or how much.

After each video, four statements were presented to understand the users' perception of the video. The four statements were:

Q1) I notice portions of this video were choppy.
Q2) The choppy portions of the video are distracting.
Q3) I notice portions of this video are blurry.
Q4) The blurry portions of the video are distracting.

The same four statements were presented after each video. A 5-point Likert scale was used to gather participant feedback after each video was shown. The degrees of the 5-point Likert scale in descending vertical order were: *strongly agree*, *somewhat agree*, *neutral*, *somewhat disagree*, *strongly disagree*. For Q2 and Q4, we also provided a *not applicable* option, since these answers depended on those to Q1 and Q3, respectively. (A participant cannot agree or disagree that the perceived choppiness or blurriness of a video was found to be distracting if choppiness or blurriness was not noticed in the first place.)

[3] The Registry of Interpreters for the Deaf, Inc. (RID) is the national licensing body for professional interpreters. The second author holds two forms of certification from RID: CI (Certificate of Interpreting) and Ed: K-12 (specialist certification in primary and secondary education).

After viewing the four videos, participants were asked background questions which included:

- What is your age?
- What is your gender?
- Do you speak ASL?
- If applicable, how many years have you spoken ASL?
- If applicable, from whom did you learn ASL?
- What language do you prefer to communicate with family?
- What language do you prefer to communicate with friends?
- Are you Deaf?
- Do you use computer instant messenger services like Skype, G-mail chat, *etc.*?
- Do you use a video phone?
- Do you use video relay services?

6. RESULTS

In the online survey, we were interested in the effects of VFR and/or VSR on video quality perception. There were 148 participants fluent in ASL (80 men, 65 women, and 3 who did not specify). Their age ranged from 18-75 years old and all but four participants were deaf. All but sixteen participants indicated that they own a cell phone and use it to text message. Finally, all but eleven participants indicated that they use video phones and use video relay services.

6.1 Analysis of Data

A nonparametric factorial analysis was used to analyze our 5-point Likert scale responses for the four questions presented after each video in our online survey. Performing a parametric F-test (ANOVA) was not appropriate because the data were not normally distributed, were ordinal in nature, and were bounded by the scale endpoints. One approach would be to use a rank transform (RT) method [13], but prior statistical research [15] has shown this to be unreliable for interaction effects (*i.e.*, for exploring VFR*VSR). Therefore, we used the *aligned rank transform (ART)* [16] procedure, which preserves interaction effects by first aligning the data [28] before ranking it. Then, a repeated measures ANOVA is performed on the aligned ranks. Readers interested in the nonparametric factorial ART procedure are directed to prior work [15,16,27,28].

In the analysis of Q2, we used 430 of 596 data points which represented responses from participants who marked 3-5 (neutral-strongly agree) in Q1, or who did not indicate Q2 was "N/A." In the analysis of Q4, we used 445 of 596 data points which represented responses from participants who marked 3-5 (neutral-strongly agree) in Q3, or who did not indicate Q4 was "N/A." Therefore, Q2 and Q4 only analyzed the responses from participants who *did* notice choppy or blurry video by marking 3-5 (neutral-strongly agree) for Q1 and Q3, respectively. Table 1 displays the mean values of applying VFR or VSR for Q1-Q4.

6.2 Perceived Choppiness (Q1 and Q2)

Q1 asked participants if they noticed choppy sections of video when VFR and/or VSR were turned on or off. Q2 followed by asking if choppy video sections were distracting. When VFR was on, participants unsurprisingly felt the video was choppier than when VFR was off (F(1,591)=80.94, *p*<.001). This result can also be seen in Table 1 where the mean value for Q1increased when VFR was turned off to on.

Table 1: Mean values of applying VFR or VSR for Q1-Q4.

	VFR	mean	standard error	VSR	mean	standard error
Q1	off	3.13	.09	off	3.57	.09
	on	4.12	.07	on	3.68	.08
Q2*	off	3.87	.08	off	4.07	.07
	on	4.25	.06	on	4.11	.07
Q3	off	3.50	.09	off	3.15	.08
	on	3.90	.07	on	4.25	.06
Q4*	off	4.11	.07	off	3.95	.08
	on	4.19	.07	on	4.30	.06

*Mean calculated from participants who marked 3-5 (neutral-strongly agree) from the previous question and did not mark "N/A."

Having VSR on or off had no effect on the perceived choppiness of the video (F(1,591)=3.58, *n.s.*). However, there was a significant VFR*VSR interaction (F(1,591)=8.09, *p*<.01). An important finding, as Figure 7 demonstrates, is that when VFR was on, the use of VSR significantly lowered the perceived choppiness of the video (F(1,591)=23.48, *p*<.001).

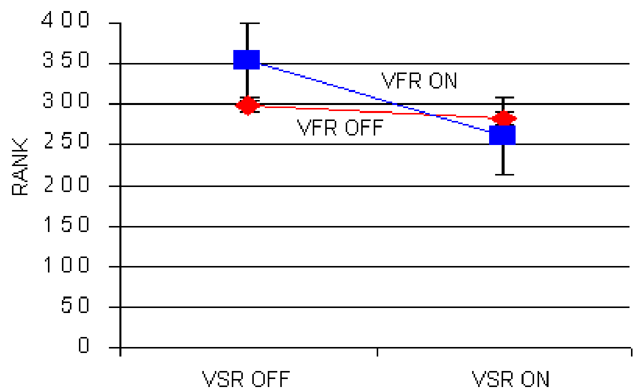

Figure 7: VFR*VSR Interaction for Q1. Note the Y-axis is the rank used by the nonparametric analysis procedure. Lower values indicate less perceived choppiness.

For Q2, participants who marked 3-5 (neutral-strongly agree) in Q1, or who did not indicate that Q2 was "N/A," we found that when VFR was on, they felt that the choppiness *was* distracting (F(1,425.3)=18.10, *p*<.001). Similar to the results found in Q1, whether VSR was on or off had no effect on participants feeling that choppiness was distracting (F(1,425)=3.86, *n.s.*). There was no VFR*VSR interaction (F(1,425)=1.65, *n.s.*).

6.3 Perceived Blurriness (Q3 and Q4)

Q3 asked participants if they noticed blurry sections of video when VFR and/or VSR were turned on or off. Q4 then asked if blurry video sections were distracting. Expectedly, when VSR was on, participants noticed the video was blurrier than when VSR was off (F(1,591)=131.57, *p*<.001). Unexpectedly, participants felt that when VFR was on, the video also appeared more blurry than when VFR was off (F(1,591)=21.95, *p*<.001). There was a significant VFR*VSR interaction (F(1,591)=18.99, *p*<.001). As Figure 8 shows, when VSR was off, whether VFR was on or off did not matter for perceived blurriness (F(1,591)=2.20, *n.s.*). But when VSR was on, the use of VFR significantly lowered the perceived blurriness of the video (F(1,591)=21.90, p<.001).

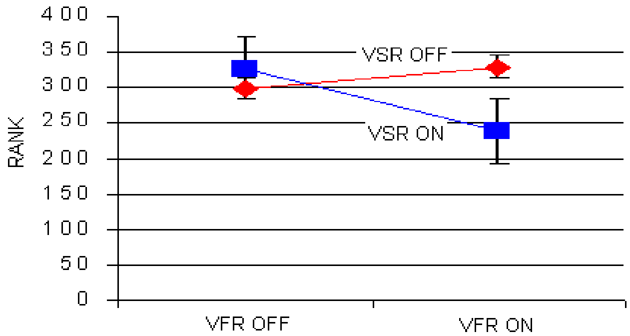

Figure 8: VFR*VSR Interaction for Q3. Note the Y-axis is the rank used by the nonparametric analysis procedure. Lower values indicate less perceived blurriness.

For Q4, participants who marked 3-5 (neutral-strongly agree) on Q3, or who did not indicate that Q4 was "N/A," we unexpectedly found that when VFR was on they perceived an increase in blurriness of the video (F(1,440.2)=7.91, *p*<.01). Not surprisingly, as Table 1 shows, when VSR was on, participants felt that the blurriness was more distracting than when VSR was off (F(1,440)=26.26, *p*<.001). Finally, there was a significant VFR*VSR interaction (F(1,440.1)=5.71, *p*<.05). As Figure 9 demonstrates, when VSR was off, whether VFR was on or off did not contribute to perceived blurriness to cause distractions (F(1,440.2)=0.34, *n.s.*) despite switching VFR off to on contributing to perceived blurriness in Q3. But when VSR was on, the use of VFR significantly reduced the distracting nature of perceived blurriness of the video (F(1,440)=9.38, p<.05).

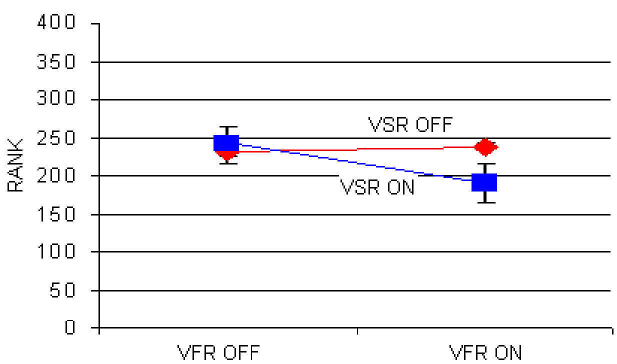

Figure 9: VFR*VSR Interaction for Q4. Note the Y-axis is the rank used by the nonparametric analysis procedure. Lower values indicate less perceived distraction due to blurriness.

7. DISCUSSION

Although one would expect to find that VFR produces perceived video choppiness and VSR produces perceived video blurriness, and indeed we found this, we also found that when *both* VFR and VSR are used, they largely ameliorate the choppiness and blurriness perceived, *i.e.*, they each improve the use of the other. A reason for this improvement could be that the blurriness caused by VSR "smoothes out" the choppy effect caused by VFR. This smoothing effect has been found in prior work to improve perception of shaky video quality when video compression is introduced [6]. It has also been found that shaky video with low temporal movement, like a home cooking show, does not degrade perceptual quality as does shaky video containing high action motion like a sports game [10]. Therefore, our findings concerning the significant VFR*VSR interactions for Q1 and Q3 indicates that VFR and VSR may work together to produce a smoothing effect. For Q3 and Q4 it was surprising to find that applying VFR increased participants' perception of blurriness since that algorithm does not objectively contribute to blurry video quality. This could be a result of participants noticing a change in video quality due to VFR and applying what they see to answer Q3 and Q4. In addition, since Q3 and Q4 were specific questions tailored to perceived blurriness, the participants' interpretation of blurriness may not have been what we intended.

8. FUTURE WORK AND CONCLUSION

In this work, we presented two new power saving algorithms, compared their performance, and found that applying either VFR or VSR alone extended the battery life to about the same battery duration. As expected, applying both VFR and VSR saved the most battery power. Conveniently, these together also produce better perceptual results than either used alone. Also, we developed a method to create a linguistically accessible online survey.

The results of the online survey suggest that the application of both VFR and VSR may be the preferred power saving algorithm, but there is more work to be done. Since our findings suggest that the application of both VFR and VSR may cause a smoothing effect, we would like to investigate whether this decreases conversational breakdowns or repeat requests among MobileASL users. A laboratory study where users engage in a conversation as opposed to watching prerecorded video with the power saving algorithms applied may reveal whether this is the case.

Applying both VFR and VSR saved the most battery power and it decreased the perceived blurriness and choppiness caused by each algorithm alone. This indicates that manipulating both the temporal *and* spatial resolution of a video to save battery power is a good approach. The most common feedback received during the online survey was, "when will MobileASL be available?" This demonstrates that there is a need for real-time mobile sign language communication and increasing the battery duration contributes to improving MobileASL technology for mainstream use.

9. ACKNOWLEDGMENTS

We thank our participants for their time. We also thank Bruce Peterson from Sorenson for sending our survey to their customers and Rahul Vanam for the variable spatial resolution idea. This work was supported by the National Science Foundation under grant IIS-0811884. Any opinions, findings, conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect those of the National Science Foundation.

10. REFERENCES

[1] Deafdigest. (2010). 2010, from http://deafdigest.com/

[2] MobileASL. University of Washington. (2010). from http://www.cs.washington.edu/research/MobileASL/

[3] QuickTime7. (2010). from http://www.apple.com/quicktime/download/

[4] Akamatsu, C. T., Mayer, C., & Farrelly, S. (2006). An Investigation of Two-Way Text Messaging Use With Deaf

Students at the Secondary Level. *Journal of Deaf Studies and Deaf Education, 11*(1), 120-131.

[5] Allen, B., Meyers, N., Sullivan, J., & Sullivan, M. (2002). American Sign Language and End-Of-Life Care: Research in the Deaf Community. *HEC Forum*, *14*(3), 197-208.

[6] Chang, Y., Carney, T., Klein, S., Messerschmitt, D., & Zakhor, A. (1998). Effects of Temporal Jitter on Video Quality: Assessment Using Psychophysical Methods. Paper presented at the *SPIE: Human Vision and Image Processing*.

[7] Cherniavsky, N. (2008). Activity Analysis of Sign Language Video for Mobile Telecommunication. Ph.D. Thesis, Computer Science and Engineering Department, University of Washington, Seattle.

[8] Cherniavsky, N., Cavender, A. C., Ladner, R. E., & Riskin, E. A. (2007). Variable Frame Rate for Low Power Mobile Sign Language Communication. Paper presented at *ASSETS'*, Tempe, Arizona, USA.

[9] Cherniavsky, N., Chon, J., Wobbrock, J. O., Ladner, R. E., & Riskin, E. A. (2009). Activity Analysis Enabling Real-Time Video Communication on Mobile Phones for Deaf Users. Paper presented at the *ACM Symposium on User Interface Software and Technology (UIST)*, Victoria, British Columbia.

[10] Claypool, M., & Tanner, J. (1999, Oct. 30- Nov. 5). The Effects of Jitter on the Perceptual Quality of Video. Paper presented at *ACM Multimedia*, Orlando, FL, USA.

[11] Coates, J., & Sutton-Spence, R. (2001). Turn-Taking Patterns in Deaf Conversation. *Journal of Social Linguistics*, 507-529.

[12] Collins, S., & Petronio, K. (1998). What Happens In Tactile ASL? In C. Lucas (Ed.), *Pinky Extension and Eye Gaze: Language Use in Deaf Communities* (pp. 18-37): Gallaudet University Press.

[13] Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician, 35*(3), 124-129.

[14] Cumming, C. E., & Rodda, M. (1989). Advocacy, Prejudice, and Role Modeling in the Deaf Community. *Journal of Social Psychology, 1*(129), 5-12.

[15] Higgins, J. J., Blair, R. C., & Tashtoush, S. (1990). The aligned rank transform procedure. *Proceedings of the Conference on Applied Statistics in Agriculture*, 185-195.

[16] Higgins, J. J., & Tashtoush, S. (1994). An aligned rank transform test for interaction. *Nonlinear World*, 1(2), 201-211.

[17] Hogg, N. M., Lomicky, C. S., & Weiner, S. F. (2008). Computer-mediated Communication & the Gallaudet community: a preliminary report. *American Annals of the Deaf, 153*(1), 89-96.

[18] Kurtnoise. (2009). Yet Another MP4 Box User Interface for Windows Users, from http://yamb.unite-video.com/index.html

[19] Lane, H. (1992). *The mask of benevolence: disabling the deaf community*. New York, NY: Alfred A. Knopf, Inc.

[20] Lucas, C., & Valli, C. (2000). *Linguistics of American Sign Language: An Introduction.* Washington, D.C.: Gallaudet University Press.

[21] Maher, J. (1996). *Seeing language in sign: the work of William C. Stokoe.* Washington, D.C.: Gallaudet University Press.

[22] Muir, L., Richardson, I., & Leaper, S. (2003). Gaze tracking and its application to video coding for sign language. Paper presented at the *Picture Coding Symposium*.

[23] Munoz-Baell, I. M., & Ruiz, T. (2000). Empowering the Deaf. Let the Deaf Be Deaf. *Journal of Epidemiology and Community Health, 1*(54), 40-44.

[24] Ong, S. C. W., & Caird, J. K. (1996). The effect of frame rate and video information redundancy on the perceptual learning of American Sign Language gestures. Paper presented at the *CHI'96: Conference companion on Human factors in computing systems*, New York, NY.

[25] Padden, C., & Humphries, T. (2005). *Inside deaf culture.* Boston: Harvard University Press.

[26] Reagan, T. (1995). A social cultural understanding of deafness: American Sign Language and the culture of deaf people. *International Journal of Intercultural Relations*, *19*(2), 239-251.

[27] Richter, S. J. (1999). Nearly exact tests in factorial experiments using the aligned rank transform. *Journal of Applied Statistics, 26*(2), 203-217.

[28] Salter, K. C., & Fawcett, R. C. (1993). The art test of interaction: A robust and powerful rank test of interaction in factorial models. *Communications in Statistics: Simulation and Computation, 22*(1), 137-153.

[29] Sperling, G., Landy, M., Cohen, Y., & Pavel, M. (1986). *Intelligible encoding of ASL image sequences at extremely low information rates*: Academic Press Professional, Inc.

[30] Tran, J. J. (2010). Power Saving Strategies for Two-Way, Real-Time Video-Enabled Cellular Phones. M.S. Thesis, Electrical Engineering Department, University of Washington, Seattle.

[31] Traum, D. R., & Hinkelman, E. A. (1992). Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*(8), 575-599.