Eccentricity Effect of Motion Silencing on Naturalistic Videos

Lark Kwon Choi[1], Lawrence K. Cormack[2], and Alan C. Bovik[1]
[1]Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, USA
[2]Department of Psychology, The University of Texas at Austin, Austin, TX, USA

***Abstract*—We study the effect of eccentricity on flicker visibility in naturalistic videos. A series of human studies were executed in two tasks ("gaze the fixation mark" and "follow the moving object") to understand how object motion can reduce the visibility of flicker distortions as a function of eccentricity, motion speed, video quality, and flicker frequency. We found that either large eccentricity or large, coherent object motion could reduce flicker visibility. When they are combined, flicker visibility significantly decreased. Flicker visibility remained noticeable even at large eccentricity when the object was static. Human study results and statistical analysis show that highly eccentric, coherent object motion can significantly silence the awareness of flicker distortions on naturalistic videos.**

***Index Terms*—Motion silencing, eccentricity, visual masking, flicker visibility, and video quality.**

I. Introduction

The perception of visual distortions has become a commonly used factor to evaluate visual quality in pervasive digital video applications. Since humans are mostly regarded as the ultimate arbiter of digital videos, understanding how humans perceive visual distortions has been an important topic for decades not only to provide satisfactory levels of quality of experience for the end user, but also to develop successful video quality assessment models [1], [2].

Digital videos can suffer from poor visual quality due to spatiotemporal distortions caused by severe compression, transmission errors, blocking, ringing, motion compensation mismatches, defocus, mosquito effects, ghosting, jerkiness, flickering, and so forth [3]. However, the mere presence of visual distortions does not have to imply quality degradation since the visibility of distortions can be strongly reduced by the perceptual phenomenon of visual masking [1]. Spatial visual masking such as luminance masking [4] and contrast masking [5] are well-modeled and play a central role in the design of perceptual image quality models [6], [7], video compression [8], and watermarking [9].

Temporal masking is not well-modeled although it has been studied to handle distortions near scene changes [10] and in the context of video compression. The perceptual effects of luminance transitions on quantization noise [8], difficulty of focusing on details on moving objects [11], an adaptive video coder to predict visibility of noise on flat areas, textures, and edges [12], the theoretical significance of visual masking on source coding [13], and a non-linear quantizer [14] have been proposed to account for temporal masking in video compression. However, scene changes are sparse, and implementations of the video coding algorithms have been largely heuristic based on anecdotal evidence.

The recently discovered "motion silencing illusion" has shown that salient changes of objects in luminance, hue, size, and shape may appear to cease when they move [15]. Motion silencing presents that the visibility of commonly occurring temporal distortions may be strongly reduced in the presence of motion. Researchers have proposed possible explanations of motion silencing [16]-[19], but the problem remains open, and the effect has only been studied on highly synthetic stimuli such as moving dots. We executed a series of human experiments to understand the motion effect of silencing on naturalistic videos wherein flicker visibility is significantly reduced by large, coherent object motions [20].

Although motion mainly contributes to the silencing effect, the phenomenon is largely a peripheral effect that does not occur at the fixation point [15]. Hence, the role of eccentricity is highly relevant, and understanding the effect of eccentricity on motion silencing in naturalistic videos is important for making the connection between the awareness of visual distortions and motion in peripheral vision. Eccentricity has been widely studied in foveation-based video compression [21]-[23] and visual quality [24], [25]; however, the effect of eccentricity on distortion visibility combined with motion has not been explicitly analyzed.

In this paper, to understand the eccentricity effect and the combined eccentricity-motion effect on flicker visibility in naturalistic videos, we generated flicker occurring at a wide range of eccentricities on naturalistic videos and conducted a series of human studies on them in a "gaze the fixation mark" task and in a separate "follow the moving object" task. We found the empirical distributions of flicker visibility on the tested videos as a function of eccentricity, object motion, video quality, and flicker frequency, and then statistically analyzed these changes. Subject's gaze was monitored by an eye tracker (faceLAB5, Seeing Machines), and object motion in the videos was predicted by an optical flow method [26].

II. Human Subjective Study

A. Source Videos

A total of eight source videos were used in the human study [20]. Source videos were 1280 × 720 or 1920 × 720 resolution at 30fps and 10 seconds. Fig. 1 shows sample frames from the test videos. The marked areas indicate the

978-1-4799-7591-4/15/$31.00 ©2015 IEEE

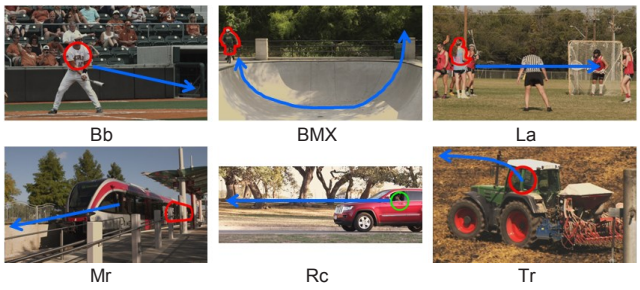

Fig. 1. Example frames of the video stimuli. The marked areas indicate the targets of the moving objects, while the arrows show the paths of object movement. The camera was fixed for the first five scenes but moved and zoomed when acquiring the "Tr" scene.

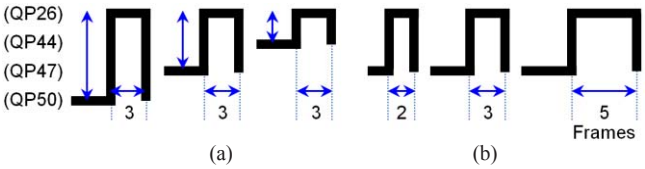

Fig. 2. Schematic illustration of the flicker distortion simulation: (a) video quality (QP) changes and (b) video frame duration changes.

targets to be judged with respect to flicker visibility on the moving objects, while the arrows show the paths of object movement. The video stimuli contain diverse object speeds. In "Bb," the motion of the batter increased from frame 150 to 208 after hitting a ball. In "BMX," a bike rider moves like a pendulum. "La" includes both static and large motions. "Mr" and "Tr" present gradual increases of object motion, where each has a different maximum speed. "Rc" contains both slow and abrupt object motion with a scene change.

B. Simulation of Flicker Distortion

We simulated quantization flicker by alternating sub-sequences of videos that were compressed by the JM H.264 encoder [27] at one of four fixed quantization parameter (QP) values: QP26, QP44, QP47, and QP50 as illustrated in Fig. 2. The flicker is caused by alternating frames from a high QP (e.g., QP44, QP47, or QP50) to QP26 and then back to low QP to high QP. Thus, the artifact is a "distortion flicker." We selected this form of flicker as such rate changes are caused by adaptive rate control algorithms. Hence, the appearance of the flicker distortion is more realistic than simple luminance flickers. To understand the effect of flicker frequency on flicker visibility as a function of eccentricity and motion, videos with QP change durations of 2, 3, and 5 frames were constructed corresponding to time-varying flicker frequencies of 3, 5, and 7.5Hz using a fixed alternating QP pair (QP47, QP26).

Unlike flicker simulations on the entire frame in [20], we generated flicker only on the small target area of the moving object. The target area is limited to within 1.5° of visual angle from the target center for all content to accurately measure the influence of eccentricity on motion silencing by minimizing variations in retinal eccentricities. We extended the QP alternating pair ranges from QP26 up to QP50 in the test videos to cover a wide range of video quality variations.

C. Test Methodology

We used a single-stimulus continuous quality evaluation (SSCQE) [28] procedure with hidden references to obtain subjective percepts of flicker visibility on the test videos. Thirty three University of Texas students served as naïve subjects. They ranged in age from 22 to 37 years old and had (corrected-to-) normal vision. Six subjects were female. We created a user interface for the study using MATLAB and the XGL toolbox [29], which interfaced with an ATI Radeon X300E graphics card in an Intel Xeon 2.93Hz CPU, 24GB RAM Windows PC. Each video was loaded into memory before its display to avoid any latency and played at the center of a 24" Dell U2410 LCD monitor (Round Rock, TX, USA) with a resolution of 1920×1080 at a 60Hz refresh rate.

Each subject participated in both of the visual tasks: Task 1, "gaze the fixation mark" and Task 2, "follow the moving object." Subjects performed Task 1 first and executed Task 2 after sufficient rest (e.g., one day). In Task 1, subjects were asked to fixate their eyes always on the fixation mark ("+" symbol) throughout the duration of video display and to rate flicker visibility on the target by moving a mouse up or down continuously. For "BMX" and "Rc," two different locations of the fixation mark were used to test eccentricity effects at the same object motion. In Task 2, subjects were asked to rate flicker visibility on the target while fixating and following the target, where the fixation mark was not used.

An instruction frame was presented before displaying each test video to indicate the fixation mark and the target. A continuously calibrated rating slider bar with Likert-like markings was shown at the right side of the screen. The rating scale ranged from 0 to 100, where the increments 0, 25, 50, 75, and 100 were marked as "Hardly," "Little," "Medium," "Highly," and "Extremely" to show the degree of perceived flicker visibility. The initial score displayed on the rating bar was "Medium" at the beginning of each video. During playback, the rating bar disappeared except for a white score gauge along the bar not to disturb video viewing. Each task lasted less than 40 minutes, and each consisting of 42 and 36 (6 hidden reference and other flicker videos) test videos in randomized order. Prior to data collection, a short training session preceded the actual test.

An eye and head tracker (faceLAB 5, Seeing Machines) was used to monitor each subject's gaze. The subjects' heads were unrestrained. Gaze was calibrated by using a 9 point calibration sequence before each task and recorded at every 1/60 s into calibrated display coordinates. The viewing distance was about 87 cm (three times the display height).

A lag response (the time difference between the perception of flicker and the movement of a mouse to rate flicker visibility) was measured for each subject. Subjects were asked to move a mouse up when a black dot on a white background flickered and to move a mouse down when the dot did not flicker. The black dot distinctively flickered for 2 or 3 s. Time duration from the start of the dot flicker on the screen to the mouse movement by the subject was measured five times, and then those values were averaged.

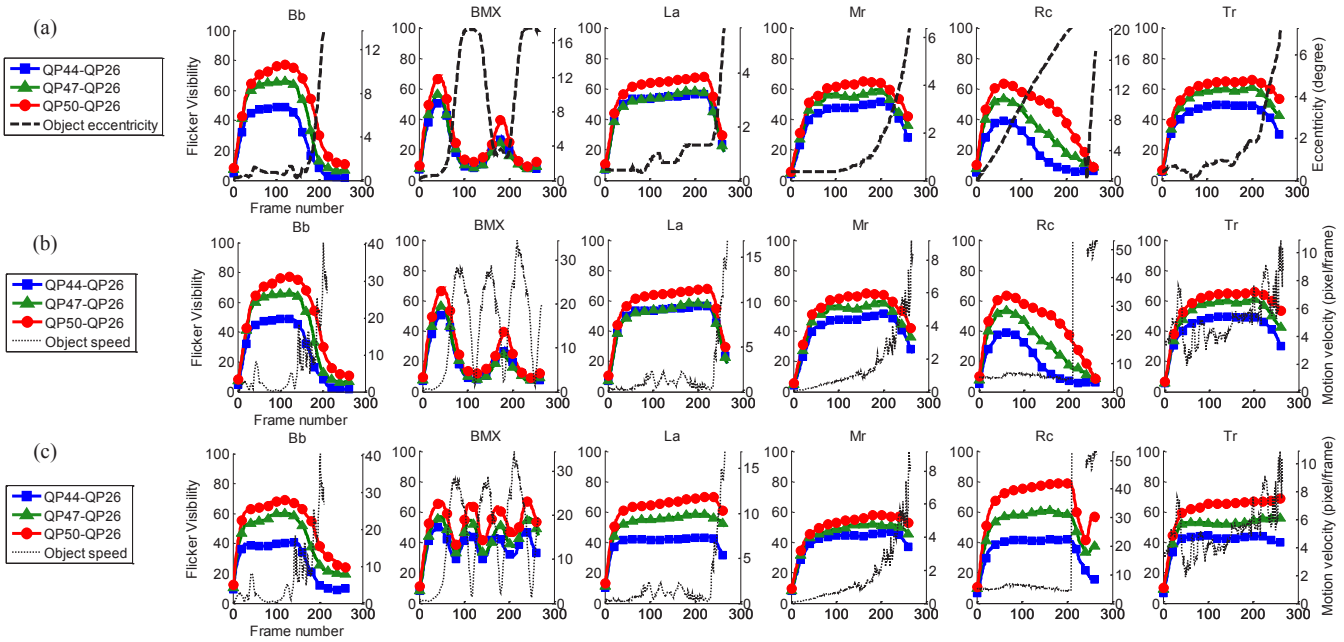

Fig. 3. Distributions of flicker visibility of the test videos for the different quality level changes at 5Hz flicker frequency. (a) Flicker visibility against object eccentricity in Task 1. (b) Flicker visibility against object motion in Task 1. (c) Flicker visibility against object motion in Task 2. Eccentricity changed differently in Task 1 as shown in (a), while eccentricity remained smaller than 1.5º for all tested videos in Task 2.

D. Processing of the Flicker Visibility Scores

We executed a subject screening procedure by examining whether subjects followed instructions for each task based on the recorded gazes. Since four subjects were rejected, the scores from the other subjects were used for the analysis.

For time synchronization between the frame at which a subject visually perceived flicker and the mechanical scoring of flicker visibility that was rated by hand, the rated score was matched after shifting the score signals by the lag response for each subject. Let s_{ijf} denote the score assigned by subject i to video j at frame f, ms_{ijf} be the visibility score of subject i to video j at frame f, and lag_i be the average lag response of subject i. Then the visibility score,

$$ms_{ijf} = s_{ij(f+lag\_i)}. \quad (1)$$

Lag responses varied from 0.57 to 1.17 s (17 - 35 frames).

We calculated the difference flicker visibility scores between the score that the subject assigned to the reference video and the scores assigned to the flicker videos in order to unbias measured flicker visibility from video content. Let ms_{ij_reff} be the flicker visibility score assigned by subject i to the reference video associated with the flicker video j after lag response matching and M_j be the total number of ratings received for video j. The difference flicker visibility score,

$$ds_{ijf} = ms_{ijf} - ms_{ij\_reff}. \quad (2)$$

The final flicker visibility score is obtained as follows,

$$fvs_{jf} = \frac{1}{M_j}\sum_i ds_{ijf}. \quad (3)$$

The flicker visibility scores range continuously from 0 to 100, where 0 means that the subject failed to or hardly perceived flicker on the target of the moving object, while 100 means that the subject perceived extreme flicker.

III. Results

A. Distribution of Flicker Visibility

The distributions of flicker visibility against eccentricity and object motion for the test videos in Task 1 and Task 2 are shown in Fig. 3. Solid lines with markers indicate the average distribution of flicker visibility for different QP alternation pairs at 5Hz flicker frequency, while the dashed and dotted lines show eccentricity and the speed of the target, respectively. Flicker visibility of the last 35 frames was not shown since flicker visibility scores were shifted to compensate for the latency of the manual responses. When targets disappeared from a scene, we omitted those intervals. The target area of the moving object associated with each video frame was manually segmented to lie inside of the 1.5º visual angle from the target center, while target moving speed was computed on successive frames using the optical flow algorithm in [27]. Eccentricity was computed using the distance from the fixation point to the target center of the moving object at a given viewing distance.

The combined effect of eccentricity and motion on flicker visibility is shown in Figs. 3(a) and 3(b). Flicker visibility substantially decreased when both eccentricity and object motion were large. Flicker visibility reduced even at smaller eccentricity than 1.5º in the presence of large motion as shown in the results of "Bb," "BMX," and "Rc" of Fig. 3(c), where flicker visibility were noticeable. However, flicker visibility was almost eliminated when both eccentricity and object motion were large as shown in Figs. 3(a) and 3(b).

Flicker visibility decreased when object motion was large or increased abruptly, while flicker visibility remained at a similar level or gradually decreased when object motion was small or steadily increased as shown in Figs. 3(b) and 3(c).

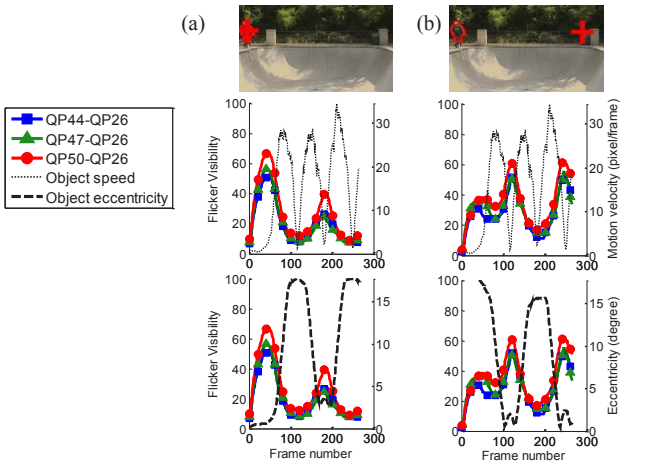

Fig. 4. Flicker visibility at the different location of gazes ("+" symbol) using the same object speed on the "BMX" test video in Task 1.

We compared flicker visibility for the same test video at two different locations of the fixation mark. The results in Fig. 4 show that flicker visibility decreases as a function of eccentricity. Moreover, we observed that flicker distortions were noticeable even at large eccentricity when the object was static (around frames 20 - 50 in Fig. 4(b)). These results demonstrate that flicker visibility on naturalistic videos strongly depends on both eccentricity and object motion.

The effect of video quality and flicker frequency was similar to the results of previous work [20]: flicker visibility decreased more when the QP difference was large (i.e., poor quality) and when the flicker frequency was large (e.g., 7.5Hz) in the presence of large motion and eccentricity, while it remained at a similar level or increased steadily when the motion was small. When eccentricity increased in the presence of small motion, flicker visibility moderately decreased as shown in "Rc" of Figs. 3(a) and 3(b).

B. Correlation Analysis

We computed the ratio of flicker visibility in Task 1 to flicker visibility in Task 2 on the same video having the same object motion but tested at different eccentricities. As shown in Fig. 5, the ratio changed little when eccentricity was smaller than approximately 1.5º but sharply decreased when eccentricity increased, which may imply that motion silencing on naturalistic videos is largely influenced by motion in foveal vision, and is mostly affected by both eccentricity and motion in peripheral vision. The results of "Rc" at different motion speeds (e.g., frames 51 - 210 and 245 - 265) for both increasing eccentricity indicate the importance of motion on flicker visibility.

We analyzed the effect of eccentricity on motion silencing of flicker by Pearson's linear correlation coefficient (LCC) for the different QP alternation pairs. Since each subject spent at least 1.67 s (50 frames, which included a judgement time and a lag response) up to 3 s to rate the first flicker visibility after a test video began, we computed LCC using the frame intervals from frame 51 to frame 265. The average magnitude of LCC for all videos was 0.8535, which implies

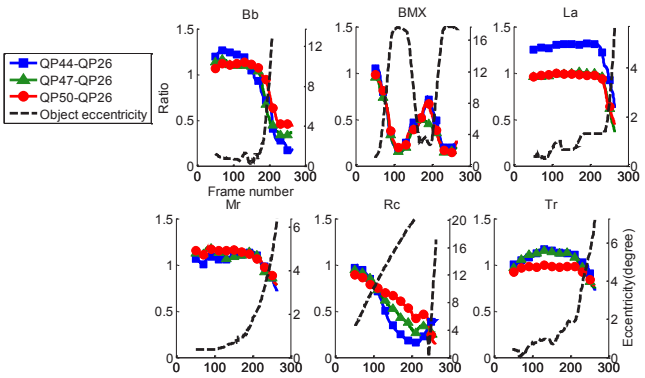

Fig. 5. Ratios of flicker visibility in Task 1 to flicker visibility in Task 2 against object eccentricity.

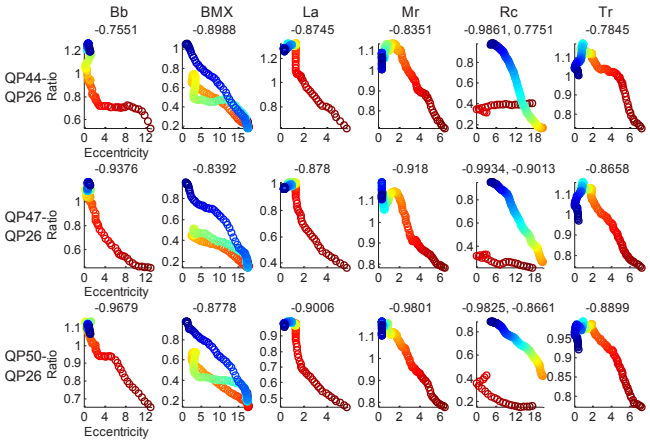

Fig. 6. Correlation coefficients between the ratios shown in Fig. 5 and object eccentricities. Time runs from cool (beginning) to hot (end). The correlation coefficient is displayed above each plot.

that eccentricity has a significant effect on flicker visibility.

C. Gaze Analysis

Subjects followed instructions well in the eye tracking environments by correctly gazing the static fixation mark in Task 1 and by following the moving targets in Task 2. The gaze traces in Task 1 were clustered near the fixation marks, while the gaze ranged around the object. The gaze ranges may have arisen in part from the eye tracker calibration errors: the average mean angular error was 0.58º and the standard deviation was 0.35º for all subjects. The gaze traces in Task 2 corresponded to the paths of the moving objects.

IV. Conclusion

We analyzed the effect of eccentricity and object motion on flicker visibility in naturalistic videos. We found that flicker visibility substantially decreased when both object motion and eccentricity were large, while flicker visibility was reduced and remained in the presence of large motion at smaller eccentricity than 1.5º. Flicker distortions were noticeable even at large eccentricity when the object was static. Based on these observations, we suggest that highly eccentric, coherent object motion can significantly silence the awareness of flicker distortions on naturalistic videos.

REFERENCES

[1] A. C. Bovik, "Automatic prediction of perceptual image and video quality," *Proc. IEEE*, vol. 101, no. 9, pp. 2008-2024, Sep. 2013.

[2] L. K. Choi, Y. Liao, and A. C. Bovik, "Video QoE metrics for the compute continuum," *IEEE Commun. Soc. Multimed. Tech. Comm. (MMTC) E-Lett.*, vol. 8, no. 5, pp. 26-29, Sep. 2013.

[3] M. Yuen and H. Wu, "A survey of hybrid MC/DPCM/DCT video coding distortions," *Signal Process.*, vol. 70, no. 3, pp. 247-278, Nov. 1998.

[4] S. E. Palmer, *Vision Science*. Cambridge, MA, USA: MIT Press, 1999.

[5] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Amer*., vol. 70, no. 12, pp. 1458-1470, Dec. 1980.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process*., vol. 13, no. 4, pp. 600-612, Apr. 2004.

[7] S. J. Daly, "Visible differences predictor: an algorithm for the assessment of image fidelity," in *Proc. SPIE Human Vis. Visual Process. and Digital Display* III, pp.2-15, 1992.

[8] A. N. Netravali and B. Prasada, "Adaptive quantization of picture signals using spatial masking," *Proc. IEEE*, vol. 65, no. 4, pp. 536-548, Apr. 1977.

[9] M. D. Swanson, B. Zhu, and A. H. Tewfik, "Multiresolution scene-based watermarking using perceptual models," *J. Sel. Areas Commun*., vol. 16, no. 4, pp. 540-550, May 1998.

[10] A. J. Seyler and Z. Budrikis, "Detail perception after scene changes in television image presentations," *IEEE Trans. Inf. Theory*, vol.11, no.1, pp.31-43, Jan. 1965.

[11] B. G. Haskell, F. W. Mounts, and J. C. Candy, "Interframe coding of videotelephone pictures," *Proc. IEEE*, vol. 60, pp. 792-800, Jul. 1972.

[12] A. Puri and R. Aravind, "Motion-compensated video with adaptive perceptual quantization," *IEEE Trans. Circuits Syst. Video Technol*., vol. 1, pp. 351-378, Dec. 1991.

[13] B. Girod, "The information theoretical significance of spatial and temporal masking in video signals," in *Proc. SPIE Human Vis. Visual Process. and Digital Display*, pp.178-187, 1989.

[14] J. D. Johnston, S. C. Knauer, K. N. Matthews, A. N. Netravali, E. D. Petajan, R. J. Safranek, and P. H. Westerink, "Adaptive non-linear quantizer," U.S. Patent 5,136,377, Aug. 4, 1992.

[15] J. W. Suchow and G. A. Alvarez, "Motion silences awareness of visual change," *Curr. Biol*., vol. 21, no. 2, pp.140-143, Jan. 2011.

[16] L. K. Choi, A. C. Bovik, and L. K. Cormack, "A flicker detector model of the motion silencing illusion," *J. Vis*., vol. 12, no. 9, pp. 777, Aug. 2012.

[17] M. Turi and D. Burr, "The motion silencing illusion results from global motion and crowding," *J. Vis*., vol. 13, no. 5, Apr. 2013.

[18] J. W. Peirce, "Is it just motion that silences awareness of other visual change?" *J. Vis*., vol. 13, no. 7, Jun. 2013.

[19] L. K. Choi, A. C. Bovik, and L. K. Cormack, "Spatiotemporal flicker detector model of motion silencing," *Perception*, vol. 43, no. 12, pp. 1286-1302, Dec. 2014.

[20] L. K. Choi, L. K. Cormack, and A. C. Bovik, "Motion silencing of flicker distortions on naturalistic videos," *Signal Process. Image Commun*., In press, Mar. 2015.

[21] Z. Wang, L. G. Lu, and A. C. Bovik, "Foveation scalable video coding with automatic fixation selection," *IEEE Trans. Image Process*., vol. 12, no. 2, pp. 243-254, Feb. 2003.

[22] L. Itti, "Automatic foveation for video compression using a neurobiological model of visual attention," *IEEE Trans. Image Process*., vol. 13, no. 10, pp. 1304-1318, Oct. 2004.

[23] Z. Chen and C. Guillemot, "Perceptually-friendly H.264/AVC video coding based on foveated just-noticeable-distortion model," *IEEE Trans. Circuits Syst. Video Technol*., vol. 20, no. 6, pp. 806-819, Jun. 2010.

[24] S. Lee, M. S. Pattichis, and A. C. Bovik, "Foveated video quality assessment," *IEEE Trans. Multimedia*, vol. 4, no. 1, pp. 129-132, Mar. 2002.

[25] A. K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," *IEEE J. Sel. Topics Signal Process*., vol. 3, no. 2, pp. 193-201, Apr. 2009.

[26] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*., pp. 2432-2439, 2010.

[27] H.264/AVC reference software coordination [Online]. Available: http://iphome.hhi.de/suehring/tml/

[28] ITU-R Rec. BT. 500-13, "Methodology for the subjective assessment of the quality of television pictures," 2012.

[29] The XGL Toolbox, 2008 [Online]. Available: http://svi.cps.utexas.edu/xgltoolbox-1.0.5.zip