TOWARD FULLY AUTOMATIC CHANNEL-MISMATCH DETECTION AND DISCOMFORT PREDICTION FOR S3D VIDEO

Alexander Bokov, Sergey Lavrushkin, Mikhail Erofeev, Dmitriy Vatolin, Alexey Fedorov

Lomonosov Moscow State University
Moscow, Russian Federation

ABSTRACT

Channel mismatch—i.e., swapped left and right views—can cause major visual discomfort when viewing stereoscopic content, but detecting it automatically remains a challenge. We present a novel method for detecting channel mismatch that significantly outperforms a prior approach when analyzing a dataset of 1,000 video clips. Moreover, this method enabled detection of 65 scenes exhibiting channel mismatch in 105 real S3D movies. Perceived discomfort can vary greatly between different scenes with channel mismatch depending on a number of factors. We conducted a study involving 59 participants to compute subjective discomfort scores for 56 scenes with channel mismatch. In this paper, we propose a model to predict these scores on the basis of scene characteristics.

Index Terms— Stereoscopic video, quality assessment, channel mismatch, discomfort prediction.

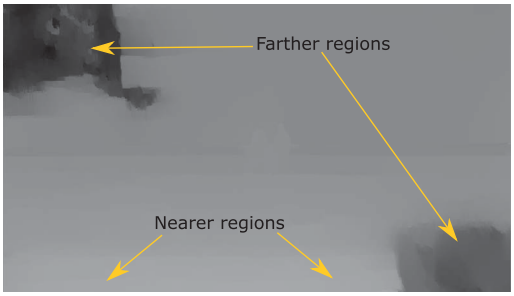
1. INTRODUCTION

The transition from 2D to stereoscopic 3D (S3D) introduces a variety of new challenges to content creators. Geometrical inconsistencies between stereoscopic views, as well as interview mismatches in color, luminance and sharpness can easily arise during S3D capture, and many of these issues demonstrably affect the perceptual quality [13]. Insufficiently accurate depth maps and poor edge processing can lead to annoying artifacts as a result of 2D-to-3D conversion [3].

Channel mismatch is a somewhat less common issue, but if present, it can lead to viewer discomfort. The left and right views can be accidentally swapped at virtually any stage of the postproduction process. Moreover, pinpointing the problem by simply watching the results in S3D can be very difficult. The level of viewer discomfort caused by channel mismatch depends on a variety of factors, including the scene's brightness, depth budget and motion characteristics. In some cases, only certain objects in a scene can be overlaid with the "eyes the wrong way around" as a result of S3D compositing errors (Figure 1).

This study was funded by RFBR, research project no. 15-01-08632 A.

(a) Source left view



(b) Disparity map

Farther Nearer

Fig. 1. Example of a partial channel mismatch involving only certain objects in a scene (the waves in the top-left and bottom-right regions of the frame). The scene is from *The Chronicles of Narnia: The Voyage of the Dawn Treader* (courtesy of Fox 2000 Pictures).

This paper presents a novel method for automatic channel-mismatch detection and visual-discomfort prediction. More precisely, our main contributions are the following:

- A channel-mismatch detection method that, as we show, outperforms a prior approach [1] (Section 3);
- A dataset containing 56 scenes with channel mismatch that we found in 105 real S3D movies using our proposed detection approach (Section 4.1);
- Results of a study involving 59 participants that evaluated the discomfort they experienced for each of the 56

978-1-5090-5743-6/16/$31.00 ©2016 IEEE

(a) Source left view

(b) Raw half-occlusion mask M_{occ}^{L}

(c) Half-occlusion confidence map C_{occ}^{L}

Fig. 2. Example of a confidence map that we use to suppress half-occlusion detection errors when applying the binocular half-occlusion criterion.

detected scenes with channel mismatch (Section 4.2);

- A model that aims to predict visual discomfort caused by channel mismatch on the basis of certain scene characteristics (Section 4.3).

2. RELATED WORK

Several recent studies have considered the problem of predicting visual discomfort caused by specific S3D artifacts. In [13], Khaustova et al. investigated how viewer annoyance depends on various technical parameters such as vertical disparity, rotation and field-of-view mismatches between the views, and color and luminance mismatches. They identified perceptual-acceptability thresholds for each of these parameters. The resulting objective metrics showed high correlation with the subjective ranks, but the authors obtained these results using a small dataset comprising three synthetic stereoscopic images. Chen et al. [5] proposed several objective metrics for luminance mismatch and evaluated their correlation with the results of a subjective experiment. They used a more diverse dataset of 30 natural stereoscopic images. In [9], the researchers investigated visual discomfort caused by temporal asynchrony.

Some researchers have proposed general methods for measuring stereoscopic-video quality that try to maximize correlation with the mean opinion score (MOS). Most of them, however, focus on transmission-related issues. Silva et al. [16] proposed a no-reference metric that takes into account overall disparity distribution, a blockiness measure and the motion characteristics of the scene. Han et al. in [10, 11] present metrics that predict perceived S3D-video quality solely on the basis of transmission parameters such as network packet-loss rate, bit rate and frame rate. The authors of [2] present a full-reference metric, having evaluated a large variety of measures, that takes into account 2D-picture quality, binocular rivalry and depth-map degradation. They used linear regression to maximize the correlation with the MOS. An optimal feature combination achieved a Spearman correlation coefficient of 0.92 on a dataset containing S3D sequences degraded by blocking and downsampling.

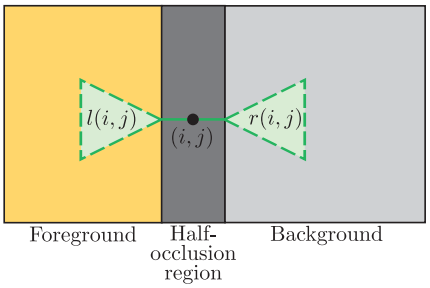


Fig. 3. Illustration of the binocular half-occlusion criterion. For any given point (i, j) inside the half-occlusion region, we determine whether the point is located on the left or right side of the foreground by using RGB histograms for the regions $l(i, j)$ and $r(i, j)$ (we assume the half-occlusion region has the same color distribution as the background).

Very few studies, however, specifically address the problems of detecting channel mismatch and evaluating the visual discomfort it introduces. M. Knee in [14] proposed a simple algorithm based on the fact that objects near the center or bottom of the screen are typically closer to the viewer than objects near the top or sides of the screen. J. Lee et al. [15] performed single-image segmentation into the foreground and background regions and compared average disparity values between them to detect channel mismatch. They evaluated their proposed approach using nine stereo sequences and showed it to be superior to subjective judgment by human subjects, who often have trouble distinguishing between correct stereoscopic images and ones with swapped left and right views. The algorithm presented in [1] uses the location of half-occlusion areas (i.e., areas that are visible only in one view of the stereopair), as well as certain characteristics of the disparity-value distribution, to detect swapped stereoscopic views. The authors validated their proposed algorithm using a dataset containing 780 random frames from 13 S3D movies. J. Bouchard et al. [4] also used an approach based on half-occlusion areas; they computed the centroids of both the left and right half-occlusion maps and compared their horizontal coordinates to detect channel mismatch. The authors used a dataset of 52 video clips to evaluate the performance of their approach. Moreover, they conducted a study with 10 human subjects and collected depth-quality scores for all 52 video clips, half of which they showed with swapped left and right views. The authors demonstrated that channel mismatch can be correctly predicted only 76% of the time when using these subjective depth-quality scores—another confirmation

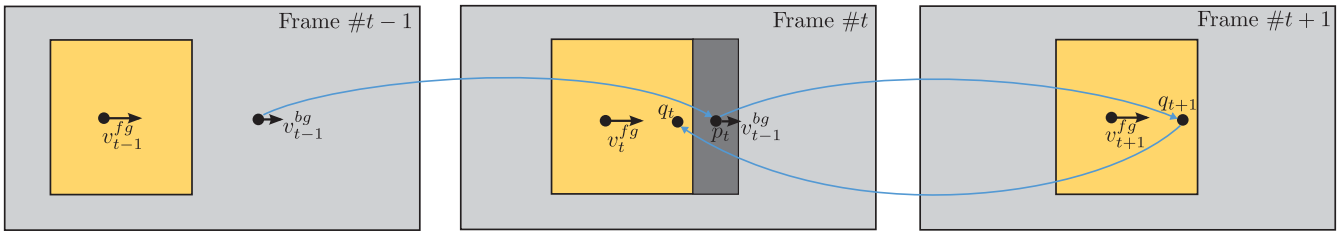


Fig. 4. Illustration of the motion-occlusion criterion. By extrapolating a background-motion vector v^{bg}_{t-1} from the previous frame for each occluded point p_t we obtain the position of the occluding point q_{t+1} in the next frame. When projected back to the original frame, p_t and q_t form the occluded/occluding point pair.

that recognizing channel mismatch is often challenging for humans.

3. PROPOSED CHANNEL-MISMATCH DETECTION METHOD

We observe that consistently detecting channel mismatch using only one or two criteria is often difficult, so we employ a weighted combination of five different criteria:

$$P_{CM} = w_p P_p + w_c P_c + w_d P_d + w_o P_o + w_m P_m, \quad (1)$$

where P_p, P_c, P_d, P_o and P_m are terms based on perspective, convexity, disparity-distribution, binocular half-occlusion and motion-occlusion criteria, respectively. We constructed each term so that positive values indicate the presence of channel mismatch, whereas negative values indicate the absence of channel mismatch. The final decision of whether a particular frame exhibits this problem depends on simple thresholding of P_{CM}.

Each criterion uses disparity maps and optical-flow fields computed by a fast local block-matching approach [17]. To filter out unreliable matches, we additionally compute the confidence values $C(i,j)$ on the basis of the LRC criterion [6], as well as local RGB variance (uniform areas are assigned a lower matching confidence). The LRC is formally defined as $|D^L(i,j) + D^R(i, j + D^L(i,j))|$, where D^L and D^R are left- and right-view disparity maps respectively. High LRC values indicate inconsistencies between left- and right-view matching results, which we assign lower confidence.

Perspective criterion: We make the simple observation that regions at the top of the frame often have higher disparity values (i.e., they are farther from the viewer) than regions at the bottom of the frame. Formally, we define the perspective-criterion term as follows:

$$P_p = \sum_{i=1}^{H-1} \min(c(i+1), c(i))(d(i+1) - d(i)). \quad (2)$$

W and H denote the input image width and height, respectively, and $d(i) = \sum_{j=1}^{W} C(i,j)D(i,j) / \sum_{j=1}^{W} C(i,j)$ is a weighted average of the disparity values $D(i,j)$ along the ith image row. The disparity-map confidence values $C(i,j)$ serve here as weights. Also, $c(i) = \sum_{j=1}^{W} C(i,j)^2 / \sum_{j=1}^{W} C(i,j)$ is a confidence value for the ith row.

Convexity criterion: The foreground often appears as a convex region surrounded by the background, which should have higher disparity values for correct stereopairs. If this assumption holds, then for each pair of points with the same disparity value, all points on the line between them should have same or lower disparity. In practice, however, the foreground is seldom strictly convex, so we assess the degree of its convexity as a sum of horizontal and vertical terms:

$$P_c = \sum_{i=1}^{H} \sum_{\substack{j_1, j_2 \in [1,W], \\ j_1 < j_2}} w(i, j_1, i, j_2) \sum_{l: j_1 < l < j_2} D(i,l) - \frac{D(i,j_1) + D(i,j_2)}{2}$$
$$+ \sum_{j=1}^{W} \sum_{\substack{i_1, i_2 \in [1,H], \\ i_1 < i_2}} w(i_1, j, i_2, j) \sum_{l: i_1 < l < i_2} D(l,j) - \frac{D(i_1,j) + D(i_2,j)}{2}. \quad (3)$$

Here, $w(i_1, j_1, i_2, j_2)$ is a weighting function defined for each point pair with coordinates (i_1, j_1) and (i_2, j_2). It gives high weights to pairs of points with similar high-confidence disparity values that are located far from each other in spatial coordinates. Formally,

$$w(i_1, j_1, i_2, j_2) = \frac{1}{1 + (D(i_1,j_1) - D(i_2,j_2))^2}$$
$$\times \sqrt{(i_2 - i_1)^2 + (j_2 - j_1)^2} \min(C(i_1,j_1), C(i_2,j_2)). \quad (4)$$

Such weights are motivated by the observation that long sequences of foreground disparity values surrounded by identical background disparity values are the strongest indicators of foreground convexity.

Disparity-distribution criterion: This criterion is based on the well-known stereoscopic principle that two-thirds of the effective disparity range should be in the positive disparity

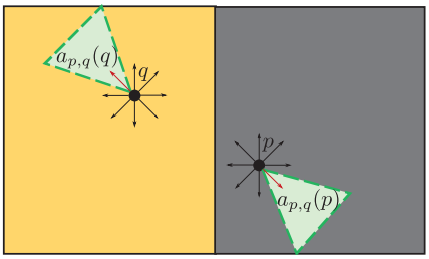


Fig. 5. For each occluded/occluding point pair (p, q), we compare the average disparity in triangular regions $a_{p,q}(p)$ and $a_{p,q}(q)$ to detect channel mismatch. We select the orientation of these regions so they face in opposite directions. All possible orientations are quantized to a set of eight directions.

space (behind the screen plane). Thus, we simply assume that most of the frame has positive rather than negative disparity values. A precise formulation of this criterion appears in [1].

Binocular half-occlusion criterion: In correct stereopairs, half-occlusion regions reside on the left side of foreground objects in the left view and on the right side in the right view. This fact enables detection of channel mismatch in scenes with sufficiently wide half-occlusion areas. To do so, we construct masks of the half-occlusion areas M^L_{occ} and M^R_{occ} for the left and right views, respectively, using the OCC method [7]. Errors in half-occlusion detection, however, can hinder the criterion's performance. We construct the half-occlusion confidence maps C^L_{occ} and C^R_{occ} that aim to filter out such errors. First, we apply a local weighted averaging to both the left- and right-view disparity maps D^L and D^R using disparity confidence values as weights. Then, horizontal derivatives of the resulting smoothed disparity maps $\hat{D}^L$ and $\hat{D}^R$ serve as half-occlusion confidence values:

$$\begin{aligned} C^L_{occ}(i,j) &= \max(-\hat{D}^L_x(i,j), 0), \\ C^R_{occ}(i,j) &= \max(\hat{D}^R_x(i,j), 0) \end{aligned} \tag{5}$$

Figure 2 shows an example of such a half-occlusion confidence map. Next, for each point in the half-occlusion region, we must determine whether it resides on the left or right side of the foreground object and then compute the final score by weighted voting among all points in the left- and right-view half-occlusion areas:

$$\begin{aligned} P_o = &\sum_{(i,j):M^L_{occ}(i,j)=1} C^L_{occ}(i,j)(H^L_{l(i,j)}[I^L(i,j)] - H^L_{r(i,j)}[I^L(i,j)]) \\ + &\sum_{(i,j):M^R_{occ}(i,j)=1} C^R_{occ}(i,j)(H^R_{r(i,j)}[I^R(i,j)] - H^R_{l(i,j)}[I^R(i,j)]), \end{aligned} \tag{6}$$

where $H^L_{l(i,j)}$ is a 16×16×16 RGB histogram computed in the region $l(i,j)$ of the left-view image I^L. Also, $l(i,j)$ and $r(i,j)$ denote triangular regions to the left and right sides of the half-occlusion area section that contains the current point (i,j) (see Figure 3). This criterion relies on the assumption that the color distribution in the half-occlusion region is more similar to that of the background than the foreground.

Movie Title	Release Year	Detected Scenes	Total Duration (sec)	Fraction of Film Duration
The Child's Eye	2010	15	57.456	0.99%
The Nutcracker in 3D	2010	9	28.946	0.45%
3D Sex and Zen: Extreme Ecstasy	2011	9	23.108	0.34%
Spy Kids 3D: Game Over	2003	5	10.302	0.20%
Hugo	2011	2	10.261	0.14%
Sharks 3D	2004	1	8.926	0.29%
Saw 3D	2010	3	6.674	0.12%
The Last Airbender	2010	2	6.590	0.11%
Dark Country	2009	4	5.756	0.11%
Creature from the Black Lagoon	1954	2	5.422	0.11%
Ghost Rider: Spirit of Vengeance	2012	1	4.630	0.08%
Stalingrad	2013	1	4.296	0.05%
Avatar	2009	1	3.336	0.03%
Spy Kids: All the Time in the World in 4D	2011	1	2.962	0.06%
Harry Potter and the Deathly Hallows: Part 2	2011	1	2.878	0.04%
The Chronicles of Narnia: The Voyage of the Dawn Treader	2010	1	2.586	0.04%
Conan the Barbarian	2011	1	1.168	0.02%
Step Up 3D	2010	1	0.876	0.01%
Clash of the Titans	2010	1	0.709	0.01%
Drive Angry	2010	1	0.668	0.01%
Bait	2012	1	0.584	0.01%
A Very Harold & Kumar 3D Christmas	2011	1	0.501	0.01%
The Three Musketeers	2011	1	0.500	0.01%

Table 1. Overall statistics of detected scenes exhibiting channel mismatch.

Motion occlusion criterion: When enough motion is present in a scene, we can also recover the relative depth ordering from occlusions in the optical-flow field. Here we compute occlusions using the OCC-Ince-Conrad method [12]. We denote the occlusion-region masks in the current frame t as $M^{t\to(t+1)}_{occ}$ and $M^{t\to(t-1)}_{occ}$, where the former marks regions that are occluded by the foreground in the next frame and the latter marks regions occluded in the previous frame. For each occluded point p_t in $M^{t\to(t+1)}_{occ}$ ($M^{t\to(t-1)}_{occ}$), we find an occluding point q_{t+1} (q_{t-1}) in the next (previous) frame under the motion-constancy assumption (see Figure 4). By projecting the occluding point back to the original frame t,

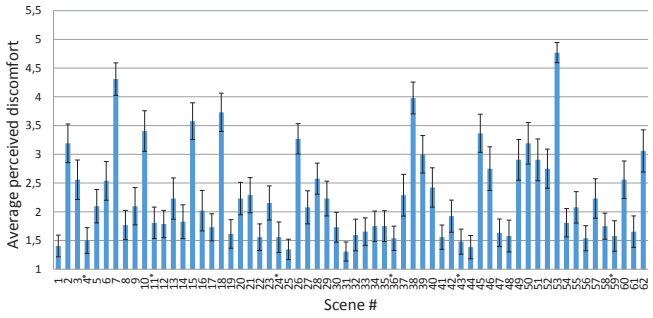


Fig. 6. Results of channel-mismatch discomfort evaluation for 62 scenes, 6 of which are control scenes with correctly ordered channels (marked with asterisks). Each value is an average of scores reported by 49 people.

we construct the set O_t of all occluded/occluding point pairs in the current frame. The criterion is based on the assumption that occluding points should have lower disparity values than the corresponding occluded points:

$$P_m = \sum_{(p,q)\in O_t} (D(a_{p,q}(p)) - D(a_{p,q}(q)))C_t(p)C_{p,q}, \quad (7)$$

where $D(a_{p,q}(p))$ denotes a weighted average of disparity values in the region $a_{p,q}(p)$. Similarly to previous criteria, we use the disparity confidence values as weights. $a_{p,q}(p)$ and $a_{p,q}(q)$ are triangular regions near the points p and q; their orientation depends on the relative position of the occluded/occluding point pair (see Figure 5). $C_t(p)$ is a confidence value of the extrapolated optical-flow vector used when locating the occluding point q. $C_{p,q}$ is an aggregated disparity confidence value for the pair (p, q); i.e.,

$$C_{p,q} = \min\left(\frac{\sum_{(i,j)\in a_{p,q}(p)} C(i,j)^2}{\sum_{(i,j)\in a_{p,q}(p)} C(i,j)}, \frac{\sum_{(i,j)\in a_{p,q}(q)} C(i,j)^2}{\sum_{(i,j)\in a_{p,q}(q)} C(i,j)}\right). \quad (8)$$

4. PROPOSED DISCOMFORT-PREDICTION METHOD

To account for the fact that perceived discomfort can vary greatly between different scenes with channel mismatch, we conduct a subjective study and propose a method for predicting discomfort owing to channel mismatch. Previous subjective studies [15, 4] have mostly focused on the human ability to recognize stereoscopic content with swapped views when it appears alongside normal stereo with correctly ordered views. We, however, aim to determine which factors contribute the most to channel-mismatch noticeability. We also aim to construct a predictive model that can be used to compute the amount of discomfort that any particular scene with channel mismatch can cause.

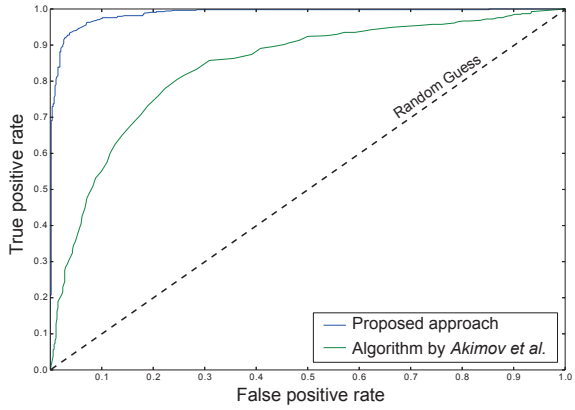


Fig. 7. ROC curves for the proposed channel-mismatch detection method and the method described in [1]. The test set consisted of 1,000 S3D video clips, half of which had swapped views.

4.1. Dataset

The effectiveness and low computational complexity of our proposed channel-mismatch detection approach has enabled us to conduct a large-scale analysis of 105 full-length S3D movies available as Blu-ray 3D discs (the present work evaluated these Blu-ray releases). We found a total of 65 scenes exhibiting manually confirmed channel mismatch in 23 different movies, for a total run time of 189 seconds. Table 1 presents more-detailed results of the analysis. Of these 65 scenes, we then used 56 (we excluded some scenes that were highly similar to others in the group) in a subjective study to evaluate perceived discomfort, and we also used them as a training set for the proposed discomfort-prediction model.

4.2. Evaluating perceived discomfort

For the study, we composed a video sequence containing all 56 selected scenes with channel mismatch. Moreover, each scene in the sequence was preceded and succeeded by scenes with the correct channel order to better simulate real viewing conditions and to provide an additional reference point for viewers. We repeated each of these three-scene fragments three times before giving the subjects time to rest and fill out the questionnaire. We asked them to rank each scene on a scale of 1 (imperceptible) to 5 (severe discomfort). The inclusion of several additional scenes with no channel mismatch ensured that the subjects ranked scenes correctly. We showed half of the subjects the sequence in reverse order to suppress the possible influence of a given scene on the subjective mark for the subsequent one. A total of 59 subjects participated in the study, but we excluded the results of 10 subjects who demonstrated maximal deviation from the statistical average or who gave high marks to the control scenes containing no channel mismatch. Figure 6 presents our results. As we ex-

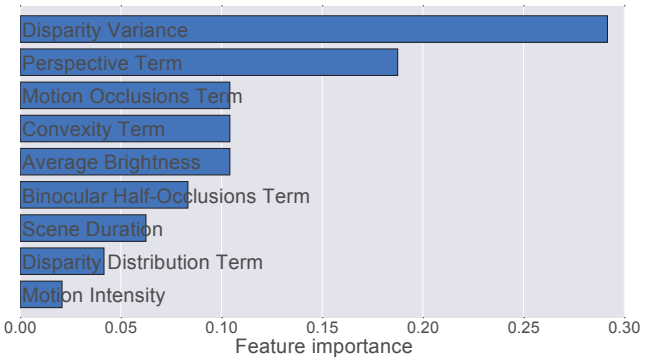


Fig. 8. Relative importance of different scene features when predicting channel-mismatch discomfort.

pected, the detected scenes differ greatly in perceived discomfort. Dark and "flat" scenes are virtually indistinguishable from scenes containing no channel mismatch (for instance, scene #31, from *The Three Musketeers*), whereas channel mismatch in some scenes (scene #53, from *Sharks 3D*) are unanimously considered to be very annoying.

4.3. Discomfort prediction

We used the results of our study to train a predictive model for channel-mismatch discomfort scores, employing the following scene features:

- Variance of disparity values;
- Average brightness;
- Motion intensity (average motion-vector length);
- Scene length;
- Values of terms P_p, P_c, P_d, P_o and P_m corresponding to the respective channel-mismatch detection criteria (see Section 3).

We tried a number of different regression methods, including linear, polynomial and random-forest-based methods. The best results, however, came from gradient boosting [8], so we used it to predict the channel-mismatch discomfort scores on the basis of the scene features listed above (we assessed the performance by using leave-one-out cross-validation).

5. EXPERIMENTAL EVALUATION

5.1. Channel-mismatch detection performance

We evaluated the performance of our proposed channel-mismatch detection approach using a dataset containing 1,000 randomly selected video clips with manually validated channel order from five recent S3D movies. Each video clip is 30 frames in length. To obtain a dataset with ground-truth values, we simply swapped the left and right views in 500 of the clips. We used another similarly constructed data set of 1,000 video clips to select the weights in Eq. 1. We performed a simple grid search while optimizing for the AUC score on the training set.

Our proposed approach significantly outperformed the method described in [1] on our test set (see Figure 7). Also, that alternative method uses a computationally expensive optical-flow algorithm, which leads to a relatively low computation speed ($\sim$138 seconds per 960×540 video clip). Our method, on the other hand, takes only $\sim$15 seconds on average to process the same video clip.

5.2. Discomfort-prediction performance

To evaluate the discomfort-prediction performance on a relatively small dataset comprising 56 scenes, we use leave-one-out cross-validation with a quadratic loss function. Gradient-tree-boosting regression with optimal hyperparameters achieves a cross-validation error of 0.43 (i.e., the mean-squared prediction error is 0.43 when predicting discomfort on a scale from 1 to 5). We also analyzed the relative importance of scene features (Figure 8).

6. CONCLUSION

In this paper, we proposed a novel channel-mismatch detection method based on a combination of five different criteria. It achieves an AUC score of 0.987 when tested on a dataset containing 1,000 S3D video clips, half which have swapped views; the method described in [1] garners a lower AUC score of 0.837 on the same dataset. Using our method, we detected 65 scenes exhibiting channel mismatch in 105 real S3D movies.

In addition, we conducted a study involving 59 human subjects to evaluate the perceived discomfort of the detected scenes on a scale of 1 to 5. We proposed a model to predict these scores on the basis of a scene's characteristics, achieving a cross-validation mean-squared error of 0.43.

7. REFERENCES

[1] D. Akimov, A. Shestov, A. Voronov, and D. Vatolin. Automatic left-right channel swap detection. In *International Conference on 3D Imaging (IC3D)*, pages 1–6, 2012.

[2] F. Battisti, M. Carli, A. Stramacci, A. Boev, and A. Gotchev. A perceptual quality metric for high-definition stereoscopic 3D video. In *SPIE/IS&T Electronic Imaging*, pages 939916–1–939916–8, 2015.

[3] A. Bokov, D. Vatolin, A. Zachesov, A. Belous, and M. Erofeev. Automatic detection of artifacts in converted S3D video. In *IS&T/SPIE Electronic Imaging*, pages 901112–1–901112–14, 2014.

[4] J. Bouchard, Y. Nazzar, and J. J. Clark. Half-occluded regions and detection of pseudoscopy. In *International Conference on 3D Vision (3DV)*, pages 215–223, 2015.

[5] J. Chen, J. Zhou, J. Sun, and A. C. Bovik. Binocular mismatch induced by luminance discrepancies on stereoscopic images. In *IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, 2014.

[6] G. Egnal, M. Mintz, and R. P. Wildes. A stereo confidence metric using single view imagery with comparison to five alternative approaches. *Image and vision computing*, 22(12):943–957, 2004.

[7] G. Egnal and R. P. Wildes. Detecting binocular half-occlusions: empirical comparisons of five approaches. *IEEE Transactions on pattern analysis and machine intelligence*, 24(8):1127–1133, 2002.

[8] J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.

[9] L. Goldmann, J.-S. Lee, and T. Ebrahimi. Temporal synchronization in stereoscopic video: Influence on quality of experience and automatic asynchrony detection. In *IEEE International Conference on Image Processing*, pages 3241–3244, 2010.

[10] Y. Han, Z. Yuan, and G.-M. Muntean. No reference objective quality metric for stereoscopic 3D video. In *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, pages 1–6, 2014.

[11] Y. Han, Z. Yuan, and G.-M. Muntean. Extended no reference objective quality metric for stereoscopic 3D video. In *IEEE International Conference on Communication Workshop (ICCW)*, pages 1729–1734, 2015.

[12] S. Ince and J. Konrad. Geometry-based estimation of occlusions from video frame pairs. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 933–936, 2005.

[13] D. Khaustova, J. Fournier, E. Wyckens, and O. Le Meur. An objective method for 3D quality prediction using visual annoyance and acceptability level. In *SPIE/IS&T Electronic Imaging*, pages 93910P–1–93910P–17, 2015.

[14] M. Knee. Getting machines to watch 3D for you. *SMPTE Motion Imaging Journal*, 121(3):52–58, 2012.

[15] J. Lee, C. Jung, C. Kim, and A. Said. Content-based pseudoscopic view detection. *Journal of Signal Processing Systems*, 68(2):261–271, 2012.

[16] A. R. Silva, M. E. V. Melgar, and M. C. Farias. A no-reference stereoscopic quality metric. In *SPIE/IS&T Electronic Imaging*, pages 93930B–1–93930B–7, 2015.

[17] K. Simonyan, S. Grishin, D. Vatolin, and D. Popov. Fast video super-resolution via classification. In *IEEE International Conference on Image Processing*, pages 349–352, 2008.