2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)

A QoE Evaluation of an Augmented Reality Procedure Assistance Application

Eoghan Hynes[1a], Ronan Flynn[1b], Brian Lee[2], Niall Murray[1a]

Department of Computer and Software Engineering[1]
Athlone Institute of Technology,
Athlone, Co. Westmeath, Ireland,
email: {e.hynes, nmurray}@research.ait.ie[a], rflynn@ait.ie[b]

Software Research Institute[2]
Athlone Institute of Technology
Athlone, Co. Westmeath, Ireland
email:blee@ait.ie

***Abstract*— Augmented reality (AR) is a key technology to enhance worker effectiveness during increasing automation of repetitive jobs in the workplace. AR will achieve this by assisting the user to successfully perform complex and frequently changing procedures. The design of AR applications for these roles is critical to their acceptability and utility. User quality of experience (QoE) will inform these design decisions. A user arrives at a quality judgment upon post-experience reflection of their degree of delight or annoyance relating to the degree of fulfilment of pragmatic and hedonic needs and expectations of the medium under consideration. QoE researchers have largely depended upon post-experience subjective reports to determine the user's QoE. Subjective reports have been shown to be biased by primacy, recency and maxima of experience stimuli. Recent research involves the identification of implicit metrics that can be used to determine user QoE continuously during a multimedia experience. This work evaluates head rotation frequency as an objective metric of user QoE.**

The literature shows that emotion is expressed in the frequency of a person's head rotation around three axes of movement (pitch, yaw and roll). Low frequency head rotation has been shown to include expression of happy emotion while high frequency exclusively expresses anger emotion. These emotions are analogous to those reflected on by a user during the quality formation process. This demo paper analyses the amount of high frequency head rotation exhibited by the user upon task completion using an AR procedure assistance application or a paper-based control. An optimal Rubik's Cube solving AR application was used as a proof of concept for AR-based procedure assistance. Preliminary results showed that the AR environment yielded higher task success rates and significantly shorter task completion durations. The AR users exhibited significantly lower amplitudes of anger frequencies in their head rotations than the control group.

Keywords*—*quality of experience; augmented reality, head-rotation frequency, affective state.

I. Introduction

As automation of repetitive procedures increases in the workplace, AR has been identified as an important technology to improve worker utility by assisting the user in performing jobs with frequently changing procedures [1]. Examples of such perpetually novel procedures in the literature include optimisation of warehouse and distribution logistics [2], machine maintenance [3] and mass customisation [4]. To this end, an optimal Rubik's Cube solver AR application was evaluated as a proof of concept for assistance with the alignment, adjustment, orientation, visual identification, inspection and verification interventions common to these novel procedures [5]. The AR application was compared against a paper-based procedure assistance medium. A significant factor that will influence the acceptability of AR for the procedure assistance use case is a positive user QoE [6]. To this end, this work evaluates user QoE of AR for procedure assistance. The traditional approach for researchers to determine user QoE has relied upon post-experience reports, which have recently been shown to be biased by primacy, recency and peak experience stimuli [7]. For this reason, recent research has focused on identifying implicit metrics of QoE such as heart rate, electrodermal activity [8]–[11], skin temperature, facial expressions [12], eye gaze [13] motion sickness, electroencephalogram [14] and gait [15]. Similarly, the focus of this work is an investigation of head rotation frequency as an objective continuous metric of QoE.

User QoE can influence the user's affective state and vice-versa [16], [17]. This includes affective state that can be expressed in head pose [18]–[21], including head rotation frequency [22]. Fig. 1 shows how head movement is considered in this work as rotation around three axes. Pitch refers to rotating the head up and down (nodding), yaw refers to turning the head from side to side (shaking) and roll refers to rotating the head to the side (tilting). This work evaluates the emotion expressed in high frequency head rotation by the AR and control test subject's upon task completion. The expression of emotion in head rotation frequency is evaluated as a proxy for the degree of delight or annoyance reflected upon by the test subjects during the quality formation process, indicating their QoE and likely quality judgement of the procedure assistance media.

II. Related Work

The authors of [1] describe AR as a key technology in maintaining good working experiences in the face of increasing automation, an environment with complex tasks that change rapidly. AR can potentially achieve this by assisting the human user in performing the frequently disparate procedures of this new and demanding work environment [1]. They explored how novice assemblers performed under AR and fixed-screen-based assembly assistance modalities in a between-subjects design

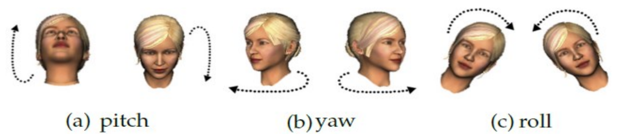

Fig. 1. Three axes of head rotation, adapted from [22].

978-1-7281-5965-2/20/$31.00 ©2020 IEEE

for a product assembly task. They not only considered objective performance metrics but also gave due consideration to the hedonic component of user experience, including user satisfaction. Their results showed that the AR users completed the assembly task quicker and reported higher satisfaction than the fixed-screen users.

The authors of [22] used head rotation frequencies and static head position as low-level features for emotion recognition. They defined the natural range of head rotation frequency as falling within 0-12Hz, with unnatural head motions above 14Hz described as occurring very rarely. They defined standardised head rotation frequency ranges as low (0-5Hz), intermediate (5-10Hz) and high (10-15Hz) and evaluated the influence of head rotation frequency on the expression of happiness, sadness, anger and neutral emotions. They concluded that low frequency head motion expresses happiness, sadness and anger together; intermediate frequencies express happiness and anger together; and high frequencies express anger exclusively. While it would be difficult to discern the component emotions in low and moderate head rotation frequencies, the high frequency range exclusively expresses anger emotion.

QoE is defined in [6] as "*the degree of delight or annoyance of a person whose experiencing involves an application, service, or system*". Delight and annoyance denote emotions [23] and head rotation has been shown to be a good metric for emotion perception [22], [24]. The hypothesis for the work presented here is that it is possible to infer the user's degree of delight or annoyance (i.e. QoE) by means of analysing the affective signals encoded in their head rotations.

Therefore, this paper proposes the use of the high frequency head rotation to evaluate a user's expression of anger as a proxy for the level of annoyance experienced during an AR-based procedure assistance application. The level of annoyance expressed by the users in a paper-based procedure assistance control medium is used as a baseline for comparison.

III. Experimental Setup

A. *AR procedure assitance modality*

An optimal Rubik's Cube solving task was used as a proof of concept for complex procedure assistance in AR and a paper-based control medium. In the AR condition, the Rubik's Cube solving instructions were presented using the Meta2 AR head mounted display (HMD). The Meta2 was a prototype AR HMD aimed at developers in which augmentations are rendered on a screen positioned in front of the wearer's eyes. The Meta2 was never released as a consumer product. The Meta2 had a 90-degree field of view; 2.5K screen resolution with a 60Hz refresh rate; a 720p front-facing RGB camera; and 9 ft (2.7 m) USB cable for video, data & power. The AR application developed in this work used the Kociemba algorithm [25] to solve the standard 3x3 Rubiks Cube in the fewest possible number of moves (optimally) from any scrambled state. Solving the cube in this way from the set all of possible starting positions (43 quintillion [26]) is "*not something which humans can do*" [27] un-assisted. At the beginning of each test, the front-facing camera on the Meta2 AR HMD was used to scan all faces of the scrambled cube. The AR application tracked the Rubik's Cube using a combination of OpenCV filters, colour and shape detection algorithms. C# was used to describe the affine features of a standard 3x3 Rubik's Cube for real-time detection from the input video feed. When the Rubik's Cube was successfully scanned at the beginning of each test in the AR condition, the AR application proceeded to heuristically solve the cube. Once solved, the AR application guided the AR user step-by-step by displaying the shortest path to the solved state directly in their field of view, one instruction at a time. The AR application was adapted from an online repository[1], originally developed for Android™[2] devices. This application was translated from Java to C# for development in Unity 3D™ using the OpenCV library.

B. *Paper-based procedure assistance modality*

Standardised testing necessitates that the procedure be the same in each test. To ensure this, each test began with the Rubik's Cube in the same initial position. The superflip position [28] was used for this, being the furthest distance from the solved state, guaranteeing the maximum of 20 moves using the optimal solution algorithm. This standardisation allowed for the creation of the paper-based control medium. The test subjects who followed the paper-based control medium instructions, the control group (CG), were provided with the same set of procedure instructions as the AR group, but printed in a 22-page A4 paper instruction manual. Each page in the manual had one printed text instruction.

C. *The assistance instructions and instruction progression*

Each instruction described the angle and direction to turn the Rubik's Cube face in question. The face names in the instruction were colour coded by referencing the face's centre tile. An example of such an instruction is:

"Rotate the face with the red at its centre 90º clockwise."

To further standardise testing, instruction progression was user-controlled in both test conditions as in [29]. The CG progressed through their instructions by turning each page of the instruction manual as required. The AR test subjects were presented with one instruction at a time in their field of view in the HMD. They progressed through the instructions, which consisted of 2D text, using keyboard input.

D. *Implicit QoE metric: head rotation frequency*

The test subject's head rotation was estimated in radians using a desk mounted Logitech 1,080p video camera and the OpenFace facial recognition application [30]. An eight second duration [31] of post-task head rotation radians was used to evaluate the test subject's emotional state at task completion. Eight seconds allows for the onset-apex-offset perception cycle of affect in head pose. Head rotation radians (rad) gathered after using the particular assistance modality limited the influence of the medium itself on head rotation. A sample rate of 27 FPS allowed for detection of frequencies up to 13.5 Hz in line with the Nyquist-Shannon sampling theorem. This allowed

[1] https://github.com/AndroidSteve/Rubik-Cube-Wizard
[2] **Android** is a **trademark** of Google LLC.

for analysis of the full range of natural head rotation frequencies defined in [22]. Linear interpolation [32] of a maximum of 2 FPS was used for up/down sampling of the time domain signal. This time domain signal was passed through a fast Fourier transform in MATLAB™ for frequency domain analysis of head rotation. The resulting radian amplitudes, divided into 22 high frequency bins from 10Hz to 13.5Hz, were used to evaluate the degree of annoyance experienced by the test subjects as expressed in these anger frequencies [22].

IV. Demonstration and Preliminary Reults

Convenience sampling led to a sample size of 48 test subjects with a mean age of 32 (σ 10). Sampling excluded those familiar to solving the Rubik's Cube. The experimental methodology consisted of five distinct phases. In the information sharing phase (1) the volunteer test subjects were informed that they would be required to solve a Rubik's Cube under one of the test conditions. Upon giving informed consent, test subjects were screened (2) for visual acuity and spatial cognition. No test subjects were excluded from testing during this screening phase. Next, the test subjects were trained (3) in using the Rubik's Cube manipulation instructions. This was followed by the practice phase (4), in which the test subject's understanding of the Cube manipulation instructions was evaluated. Upon successful demonstration of understanding, the test subjects proceeded to the testing phase (5) under one of the test conditions. In testing, the test subjects were presented with a Rubik's Cube in the superflip position, requiring twenty manipulations to solve correctly.

The AR group proceeded to testing wearing the AR HMD while the control group were presented with a paper-based instruction manual. Instructions consisted of 2D text to standardise both testing environments. Each instruction described the direction and amount of rotation to apply to each Cube face. There were three fundamental instructions used to solve the Rubik's Cube. These were: (a) a 90º clockwise, (b) a 90º anti-clockwise, or (c) a 180º clockwise rotation, of the face in question. The face in question is identified by referring to the colour of its central tile, as this never changes location on the face. The test subjects firstly had to locate the correct Rubik's Cube face by rotating the entire Rubik's Cube in theirs hands. Upon locating the correct face, the test subjects then had to rotate the individual Rubik's Cube face by the required amount in the correct direction. After the twentieth instruction, a final statement reading "The Cube should now be solved" was presented in both test environments, informing the test subject that the test was complete. If the test subjects correctly followed all twenty instructions, they finished the test with a correctly solved Rubik's Cube.

Preliminary results showed that AR and paper-based instruction yielded successful procedure completion in 96% and 92% of tests respectively (p=0.555, α=0.05). Mean task completion times were 02:24s and 02:47s (p=0.033, α=0.05) for AR and paper-based instruction respectively. The control group consistently expressed significantly higher (p=0.000, α=0.05) radian amplitudes of the high range frequencies on each axis of head rotation during the post task sample. These results are shown in Fig. 2, which shows each groups minimum, mean and maximum amplitudes for pitch yaw and roll combined, plotted across each of the 22 high frequency bins (from 10Hz to 13.5Hz). The CG had higher maximum amplitudes of rotation on all three axes with pitch = 0.48 rad, yaw = 0.20 rad and roll = 0.27 rad greater than the AR group. Pitch, yaw and roll accounted for 41%, 35% and 24% of the CG's higher mean amplitudes respectively. Fig. 3 shows each individual's mean amplitude of the high range frequencies plotted against their task duration. Linear regression shows that an increase in task duration accounted for a minor increase in the amplitudes of high range head rotation frequencies in both test groups with R^2=0.005 for pitch (blue), R^2=0.008 for yaw (yellow) and R^2=0.003 for roll (green). Head rotation estimation confidence values were 91.4% and 98.0% in the AR group and CG respectively, where the AR HMD resulted in a loss of an average of 6.6% estimation confidence due to the partial occlusion of the AR user's upper head.

V. Conclusion

This work investigated the use of head rotation frequency as an indicator of QoE, with a focus on AR. AR yielded utility gains over the paper-based control by improved task success rates and significantly reduced task durations. While the paper-based control group performed worse than the AR group they also expressed circa twice the mean amplitudes of anger emotion in their head rotation frequencies during a post task sample. Future work proposes to use head rotation frequency to evaluate user QoE continuously during usage. The influence of head rotation estimation confidence on these results will be investigated. The proposed future evaluation system allows for cross correlation between head rotation frequency and established subjective QoE metrics.

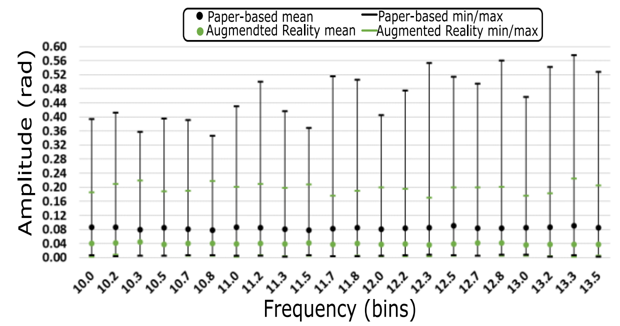

Fig. 2. Group minimum, mean and maximum amplitutdes of combined high range pitch, yaw and roll frequencies.

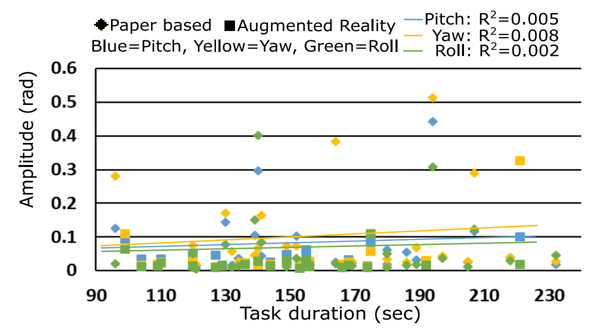

Fig. 3. Individual mean amplitudes of high range pitch yaw and roll frequencies plotted against task duration.

REFERENCES

[1] S. Egger-Lampl, C. Gerdenitsch, L. Deinhard, R. Schatz, and P. Hold, ‘Assembly Instructions with AR: Towards measuring Interactive Assistance Experience in an Industry 4.0 Context’, in 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), Jun. 2019, pp. 1–3, doi: 10.1109/QoMEX.2019.8743266.

[2] B. Mueck, M. Höwer, W. Franke, and W. Dangelmaier, ‘Augmented Reality applications for Warehouse Logistics’, in Soft Computing as Transdisciplinary Science and Technology, vol. 29, A. Abraham, Y. Dote, T. Furuhashi, M. Köppen, A. Ohuchi, and Y. Ohsawa, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 1053–1062.

[3] R. Schlagowski, L. Merkel, and C. Meitinger, ‘Design of an assistant system for industrial maintenance tasks and implementation of a prototype using augmented reality’, in 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Dec. 2017, pp. 294–298, doi: 10.1109/IEEM.2017.8289899.

[4] F. Loch, F. Quint, and I. Brishtel, ‘Comparing Video and Augmented Reality Assistance in Manual Assembly’, in 2016 12th International Conference on Intelligent Environments (IE), Sep. 2016, pp. 147–150, doi: 10.1109/IE.2016.31.

[5] U. Neumann and A. Majoros, ‘Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance’, in Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180), Mar. 1998, pp. 4–11, doi: 10.1109/VRAIS.1998.658416.

[6] S. Möller and A. Raake, Quality of Experience, Advanced Concepts, Applications and Methods. Springer, 2013.

[7] S. S. Sabet, C. Griwodz, and S. Möller, ‘Influence of primacy, recency and peak effects on the game experience questionnaire’, in Proceedings of the 11th ACM Workshop on Immersive Mixed and Virtual Environment Systems, Amherst, Massachusetts, Jun. 2019, pp. 22–27, doi: 10.1145/3304113.3326113.

[8] D. Egan, S. Brennan, J. Barrett, Y. Qiao, C. Timmerer, and N. Murray, ‘An evaluation of Heart Rate and ElectroDermal Activity as an objective QoE evaluation method for immersive virtual reality environments’, in 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX), Jun. 2016, pp. 1–6, doi: 10.1109/QoMEX.2016.7498964.

[9] C. Keighrey, R. Flynn, S. Murray, and N. Murray, ‘A Physiology-based QoE Comparison of Interactive Augmented Reality, Virtual Reality and Tablet-based Applications’, IEEE Trans. Multimed., pp. 1–1, 2020, doi: 10.1109/TMM.2020.2982046.

[10] C. Keighrey, R. Flynn, S. Murray, and N. Murray, ‘A QoE evaluation of immersive augmented and virtual reality speech language assessment applications’, in 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), May 2017, pp. 1–6, doi: 10.1109/QoMEX.2017.7965656.

[11] C. Keighrey, R. Flynn, S. Murray, S. Brennan, and N. Murray, ‘Comparing user QoE via physiological and interaction measurements of immersive AR and VR speech language therapy applications.’, 25th ACM Int. Conf. Multimed. ACM MM 2017, vol. Thematic Workshop.

[12] E. Hynes, R. Flynn, B. Lee, and N. Murray, ‘A Quality of Experience Evaluation Comparing Augmented Reality and Paper Based Instruction for Complex Task Assistance’, in 2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP), Sep. 2019, pp. 1–6, doi: 10.1109/MMSP.2019.8901705.

[13] D. Concannon, R. Flynn, and N. Murray, ‘A quality of experience evaluation system and research challenges for networked virtual reality-based teleoperation applications’, in Proceedings of the 11th ACM Workshop on Immersive Mixed and Virtual Environment Systems, Amherst, Massachusetts, Jun. 2019, pp. 10–12, doi: 10.1145/3304113.3326119.

[14] D. P. Salgado et al., ‘A QoE assessment method based on EDA, heart rate and EEG of a virtual reality assistive technology system’, in Proceedings of the 9th ACM Multimedia Systems Conference on - MMSys ’18, Amsterdam, Netherlands, 2018, pp. 517–520, doi: 10.1145/3204949.3208118.

[15] T. B. Rodrigues, C. Ó. Catháin, D. Devine, K. Moran, N. E. O’Connor, and N. Murray, ‘An evaluation of a 3D multimodal marker-less motion analysis system’, in Proceedings of the 10th ACM Multimedia Systems Conference, Amherst, Massachusetts, Jun. 2019, pp. 213–221, doi: 10.1145/3304109.3306236.

[16] ‘Context in Emotion Perception - Lisa Feldman Barrett, Batja Mesquita, Maria Gendron, 2011’. https://journals.sagepub.com/doi/abs/10.1177/0963721411422522 (accessed Feb. 14, 2020).

[17] P. Reichl et al., ‘Towards a comprehensive framework for QOE and user behavior modelling’, in 2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX), Pylos-Nestoras, May 2015, pp. 1–6, doi: 10.1109/QoMEX.2015.7148138.

[18] J. O’Dwyer, ‘Speech, Head, and Eye-based Cues for Continuous Affect Prediction’, in 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), Sep. 2019, pp. 16–20, doi: 10.1109/ACIIW.2019.8925042.

[19] J. O’Dwyer, N. Murray, and R. Flynn, ‘Eye-based Continuous Affect Prediction’, in 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), Sep. 2019, pp. 137–143, doi: 10.1109/ACII.2019.8925470.

[20] J. O’Dwyer, R. Flynn, and N. Murray, ‘Continuous affect prediction using eye gaze and speech’, in 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Nov. 2017, pp. 2001–2007, doi: 10.1109/BIBM.2017.8217968.

[21] J. O’Dwyer, R. Flynn, and N. Murray, ‘Continuous affect prediction using eye gaze’, in 2017 28th Irish Signals and Systems Conference (ISSC), Jun. 2017, pp. 1–6, doi: 10.1109/ISSC.2017.7983611.

[22] Y. Ding, L. Shi, and Z. Deng, ‘Low-level Characterization of Expressive Head Motion through Frequency Domain Analysis’, IEEE Trans. Affect. Comput., pp. 1–1, 2018, doi: 10.1109/TAFFC.2018.2805892.

[23] J. A. Russell, ‘A circumplex model of affect’, J. Pers. Soc. Psychol., vol. 39, no. 6, pp. 1161–1178, Dec. 1980, doi: http://dx.doi.org.ezproxy.ait.ie/10.1037/h0077714.

[24] A. Adams, M. Mahmoud, T. Baltrusaitis, and P. Robinson, ‘Decoupling facial expressions and head motions in complex emotions’, in 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), Xi’an, China, Sep. 2015, pp. 274–280, doi: 10.1109/ACII.2015.7344583.

[25] T. Rokicki, H. Kociemba, M. Davidson, and J. Dethridge, ‘The Diameter of the Rubik’s Cube Group Is Twenty’, SIAM Rev., vol. 56, no. 4, pp. 645–670, Jan. 2014, doi: 10.1137/140973499.

[26] ‘Solving the Rubik’s Cube Optimally is NP-complete’, ResearchGate. https://www.researchgate.net/publication/317732228_Solving_the_Rubik's_Cube_Optimally_is_NP-complete (accessed Jul. 18, 2018).

[27] T. Rokicki, ‘Why It’s Almost Impossible to Solve a Rubik’s Cube in Under 3 Seconds’, May 24, 2019. https://youtu.be/SUopbexPk3A?t=738 (accessed Jul. 09, 2019).

[28] J. Pride, ‘From Superflip to Solved’, Academic Excellence Showcase Schedule, Jun. 01, 2017. https://digitalcommons.wou.edu/aes_event/2017/all/226.

[29] S. J. Henderson and S. K. Feiner, ‘Augmented reality in the psychomotor phase of a procedural task’, in 2011 10th IEEE International Symposium on Mixed and Augmented Reality, Oct. 2011, pp. 191–200, doi: 10.1109/ISMAR.2011.6092386.

[30] T. Baltrušaitis, P. Robinson, and L. P. Morency, ‘OpenFace: An open source facial behavior analysis toolkit’, in 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Mar. 2016, pp. 1–10, doi: 10.1109/WACV.2016.7477553.

[31] M. Valstar et al., ‘AVEC 2016 - Depression, Mood, and Emotion Recognition Workshop and Challenge’, ArXiv160501600 Cs, May 2016, Accessed: Aug. 26, 2019. [Online]. Available: http://arxiv.org/abs/1605.01600.

[32] T. Blu, P. Thevenaz, and M. Unser, ‘Linear interpolation revitalized’, IEEE Trans. Image Process., vol. 13, no. 5, pp. 710–719, May 2004, doi: 10.1109/TIP.2004.826093.