Multimodal Vision–Haptic Perception of Digital Watermarks Embedded in 3-D Meshes

Nizar Sakr, *Student Member, IEEE*, Nicolas D. Georganas, *Fellow, IEEE*, Jiying Zhao, *Member, IEEE*, and Emil M. Petriu, *Fellow, IEEE*

***Abstract*—In this paper, a study is conducted to investigate the role of multisensory feedback in the perception of a watermark embedded in a haptic-enabled 3-D virtual surface. Specifically, we focus on whether the use of a bimodal approach, where a subject is looking at, and touching a 3-D mesh could in fact improve the perceptibility of a watermark as opposed to when only a single modality (vision-alone or haptic-alone) is available. The experiments are performed using a vision–haptic interface that enables users to see and touch virtual objects at the same location in space. This approach enables superior integration of vision and touch than a conventional 2-D screen-based display. An analysis of variance (ANOVA) is also conducted to statistically explore the impact of the considered modalities (vision, haptic, and vision–haptic) on the measured watermark detection thresholds across different resolutions of the underlying virtual 3-D mesh. Overall, the results suggest that relying on bimodal vision–haptic feedback is better than any of the single modalities when detecting a watermark embedded in a 3-D model. In addition, it has been assessed that the impact of the selected modality on the perceptibility of the 3-D watermark is independent of the chosen surface resolution.**

***Index Terms*—Bimodal, detection threshold, haptics, vision, watermarking.**

I. Introduction

THE increased processing capabilities of today's modern computers have contributed to the widespread use of 3-D models in several fields, such as in entertainment (video games, and movies), the medical industry (3-D visualization of medical data, education, and training), and the Internet (3-D virtual worlds and e-commerce). In turn, the protection of 3-D models from theft or tampering has received much attention in the literature [1]. An effective measure that overcomes the limitations of traditional encryption is digital watermarking, i.e., a research field that deals with the process of embedding information into digital data in an inconspicuous manner to identify the origin, owner, use, rights, integrity, or destinations of multimedia content (e.g., digital images, video, audio, and 3-D models) [1]–[5]. The first requirement of digital watermarking techniques regardless of the addressed media or application is imperceptibility [4], [6], [7]. Imperceptibility refers to the perceptual similarity between the original and watermarked data. Ideally, the perceptual quality of the watermarked media must be identical to the original. Recently, considerable progress has been made in 3-D watermarking, where the main focus has been on triangle meshes, which is the most common digital representation of 3-D models due to its simplicity and usability. Existing watermarking techniques concerning 3-D meshes host the watermark either by modifying the geometry or the connectivity of the surface (spatial domain) or by modifying some kind of spectral-like coefficients (spectral domain) [1]. Accordingly, the watermark's intrusiveness can be evaluated in terms of its visibility in the rendered version of the mesh. The evaluation of 3-D watermarking algorithms against the imperceptibility constraint has been thus far exclusively based on the sensitivity of human vision to distortion. Moreover, currently available perceptual metrics generally used to assess the quality of watermarked 3-D meshes have been validated solely through psychovisual experiments [6], [7].

In recent years, the integration of haptics into immersive virtual environments to enable sensing and manipulation of virtual 3-D objects through touch has become increasingly popular due to its many potential applications, including medical training, physical rehabilitation, and entertainment [8]–[10]. In consequence, the recent advancements of haptic technology have raised several important questions in the 3-D watermarking research community: Is the haptic sensory channel more sensitive than the visual sensory channel in detecting a watermark embedded in a haptic-enabled 3-D object? Do watermarks inspected using multimodal feedback (vision + haptic) result in very different detection thresholds from those detected using a single sensory modality (touch-only or vision-only); Or more importantly, does visual feedback, when presented together with haptic feedback, improve the perception of a watermark embedded in a 3-D mesh? If it is revealed that the lowest watermark detection thresholds are obtained under haptic-alone and/or haptic + vision conditions, as compared with vision-alone, then 3-D watermarking researchers must carefully design (or redesign) their algorithms to ensure watermark imperceptibility when both visual and haptic feedback are considered.

In this paper, a study is conducted to examine the role of multisensory feedback in the perception of a watermark embedded in a haptic-enabled 3-D mesh. Specifically, we focus on whether the use of a bimodal approach, where a subject is looking at, and

Manuscript received June 29, 2009; revised December 1, 2009. First published March 15, 2010; current version published April 7, 2010. The Associate Editor coordinating the review process for this paper was Dr. Serge Demidenko.

N. Sakr, J. Zhao, and E. M. Petriu are with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: nsakr@discover.uottawa.ca; jyzhao@discover.uottawa.ca; petriu@discover.uottawa.ca).

N. D. Georganas is with IMDEA Networks, University Carlos III of Madrid, 28911 Madrid, Spain, on leave from the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: georganas@discover.uottawa.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2010.2040970

touching a 3-D mesh could in fact improve the perceptibility of a watermark as opposed to when only a single modality (vision-alone or haptic-alone) is available. The experiments are performed using a vision–haptic interface that enables users to see and touch virtual objects at the same location in space. This approach enables superior integration of vision and touch than a conventional 2-D screen-based display. Moreover, the advantages of the use of a vision–haptic interface stem from the fact that human perception is generally multimodal, i.e., the senses of touch and vision do not independently operate but are in fact closely coupled. A statistical analysis to test for significance between modalities is also presented. An analysis of variance (ANOVA) is conducted to statistically explore the impact of the considered modalities (vision, haptic, and vision–haptic) on the measured watermark detection thresholds across different resolutions of the underlying virtual 3-D mesh.

The rest of this paper is organized as follows: In Section II, the related work is presented. In Section III, the importance of multimodal vision–haptic watermarking is emphasized. In Section IV, the materials and methods of the conducted experiments are described in detail. In Section V, the experimental results are presented and discussed. Finally, conclusive remarks are outlined in Section VI.

II. Related Work

In this section, previous works closely related to this paper are described in two parts. First, an overview is presented of similar experimental studies that attempt to assess whether multimodal conditions (haptic + vision) can produce very distinct percepts from those produced when a single sensory modality is available. The focus of this discussion is however limited to studies that explore the impact of multisensory vision–haptic feedback on the perception of surface roughness. This is due to the fact that surface distortions produced from the 3-D watermarking process can be regarded as a particular form of surface roughness. The second part of this section discusses existing work in the literature that has investigated the area of haptic digital watermarking.

A. Multimodal Vision–Haptic Perception

A fair amount of research has been devoted to understanding how observers combine multisensory (vision–haptic) inputs to attain an integrated percept. For example, whether multisensory vision + haptic conditions can produce different percepts from those produced when a single sensory modality is present, and whether the observers' behavior is consistent across different perceptual tasks, e.g., visual–haptic stiffness [11], [12], size [12], and shape [13]. Of more particular interest, a number of studies have also investigated the role of multisensory feedback on the perception of surface roughness [14]–[20].

Earlier studies of roughness perception with real surfaces have shown that multimodal conditions can lead to very different percepts than when only single modalities are considered. For example, in [16], participants were requested to judge the smoothness of real surfaces while relying on visual-only, touch-only, and visual + touch sensing conditions. It was observed that the roughness discrimination task was better performed using touch-only than vision-only, particularly for very smooth surfaces. More importantly, bimodal vision–touch perception resulted in better discrimination results than any of the single modalities alone. In [17], Lederman and Abbott suggested that the extent to which one modality can dominate another varies depending on the structure of the sensing task. It is mentioned that when participants were requested to judge surface roughness in discrepant modalities, approximately equal weighting was assigned to the visual and haptic sensory inputs. In follow-up work [18], the authors modified the instructions of the experimental task given to participants and determined that visual feedback dominated haptic when observers were asked to judge spatial density; however, haptic feedback dominated vision when observers were asked to judge roughness. In a recent paper, however [19], the authors reinvestigated this topic using virtual-reality techniques and a haptic force-feedback device. It was found that in contrast to [18], performance for roughness and density judgments did not differ significantly, i.e., the roughness estimates could not be discriminated from those of spatial density. It was also observed that for certain textures, there was evidence of vision–haptic intermodal integration that resulted in improved texture discrimination performance.

Conversely, in [20], Guest and Spence examined the role of multisensory integration in vision-tactile perception of texture. Their experiments revealed that vision and touch act as independent sources of roughness information. Consequently, there was no evidence of multimodal vision–haptic integration as roughness discrimination performance did not improve in bimodal conditions.

In [15], Poling *et al.* conducted a study to investigate whether roughness discrimination of virtual surfaces is different under bimodal visual + haptic conditions as compared with unisensory conditions where only vision or haptic feedback is available. The virtual surfaces were sinusoidal gratings that varied in amplitude and spatial period. It was determined that when the grating amplitude was very small, the bimodal (vision + haptic) condition was highly dominated by haptic feedback (i.e., performance in the vision + haptic condition is identical to that in the haptic only condition). Conversely, when the grating amplitude was large, the bimodal (vision + haptic) condition was highly dominated by visual feedback. For intermediate grating amplitudes, however, the bimodal condition proved to be better than haptic-alone and vision-alone.

B. Haptic Watermarking

Despite its importance, only one other research group has investigated watermarking of haptic-enabled 3-D models. In [2] and [21], Prattichizzo *et al.* introduced a haptic watermarking technique where a sinusoidal watermark is superimposed onto a sinusoidal host surface texture signal. The authors achieved imperceptible haptic watermarking while relying on past experimental research on human detection thresholds for sinusoidal stimuli, which depict the minimum signal strength required for producing a sensation. Prattichizzo *et al.* also demonstrated how the haptically imperceptible sinusoidal watermark can be detected by means of spectral analysis.

In [2] and [22], psychophysical experiments are conducted to measure the perceptibility of a watermark embedded in a 3-D mesh. The watermark embedding is achieved by altering the positions of the mesh vertices. The experiments were performed using several 3-D meshes with different resolutions in an attempt to establish a relationship between the watermark strength and the size of the triangular mesh components. The watermark perceptibility is inspected through a visual or a haptic interface, i.e., the user can only see or feel the virtual object. Experiments in which the subject can inspect the watermark perceptibility while simultaneously looking at and touching the virtual surface (bimodal conditions) were never demonstrated. An analysis using advanced statistical methods to test for significance between modalities (and possibly surface resolutions) was never provided. Furthermore, the experiments performed in the study relied on a small sample of participants (five subjects). This is an important limitation as it makes it more difficult to test hypotheses of differences among factors or to make any statistical generalizations.

III. Why Multimodal Vision–Haptic Watermarking?

It is intuitive to assume that a multimodal presentation of stimuli (e.g., for watermark detection) should lead to an improvement in performance when compared with the unimodal alternative. However, studies discussed in Section II-A have revealed that the role of multisensory feedback in roughness perception is not well understood (i.e., evidence for multisensory integration is mixed in this area) and certainly not straightforward as it can vary across different experimental conditions, including type of surface (real or virtual), haptic device used (tactile or force feedback), surface parameters (spatial density, roughness), etc. Specifically, in certain cases, it has experimentally been illustrated that the addition of a second modality could divide attentional resources and consequently reduce the information available from each of the individual modalities [20]. It is also suggested that this lack of multisensory integration can be due to the fact that information available from distinct modalities is possibly redundant. This finding has not only been observed in roughness perception but also in other studies investigating the role of multimodal integration, such as vision–haptic length perception [23]. Other studies [12], [15] support the notion that, in certain scenarios, an observer attends to cues from the dominant modality and discards information from the less dominant modality. Conversely, other findings in vision–haptic roughness perception [16], [19] suggest that inputs from the two modalities are efficiently integrated to produce a percept that is better than that provided by either modality alone.

Based on the various conclusions previously depicted regarding the integration of multisensory feedback in roughness perception (and assuming that 3-D haptic watermark detection is a subcategory of virtual surface roughness perception), it is impossible to predict with certainty the role of bimodal vision–haptic feedback in the detection of a 3-D watermark. To further emphasize the significance of analyzing multimodal vision–haptic watermarking, it is important to realize that the current vision-only-based imperceptibility criteria (in mesh watermarking) will ultimately be redefined if a scenario is uncovered in which watermark detection is best achieved using bimodal feedback. Consequently, 3-D watermarking researchers will be required to carefully design (or redesign) their algorithms to ensure watermark imperceptibility in multimodal vision–haptic conditions. Conversely, if it is revealed that watermark detection performance in the haptic + vision condition is identical to (or worse than) that in the haptic or vision-alone condition, then it would be confirmed that the evaluation of a watermarking algorithm in a bimodal setup is not necessary for it to be categorized as *imperceptible*.

In this paper, a study is conducted to investigate the role of bimodal vision–haptic feedback in the perception of a watermark embedded in a haptic-enabled 3-D virtual surface. To the best of the authors' knowledge, this is the first work in the literature that examines multimodal vision–haptic watermarking. The experiments are performed using a vision–haptic interface that enables users to see and touch virtual objects at the same location in space. Hence, user interactions are more natural and thus provide the ideal environment for high-precision vision–haptic experiments. Specifically, a two-way within-subject ANOVA is conducted to statistically explore the impact of the considered modalities on the measured watermark detection thresholds across different resolutions of the underlying virtual 3-D mesh. The contributions of this paper are as follows: Experimentally and statistically analyze the following: 1) whether haptic-alone is superior to vision-alone in detecting a watermark embedded in a 3-D mesh; 2) more importantly, if bimodal vision–haptic conditions can improve the watermark detectability as opposed to either modality alone; 3) if the impact of the selected modality (or modalities) on the perceptibility of the 3-D watermark is dependent on the chosen surface resolution; and 4) whether in bimodal conditions the detectability of the watermark increases or decreases with the resolution of the surface mesh.

IV. Materials and Methods

As previously mentioned, the objective of the undertaken experiments is to measure the perceptibility of a watermark embedded in a 3-D mesh while looking at, touching, or simultaneously looking at and touching the virtual surface. To simplify the analysis of the experiments, the virtual object is selected as in [2], and it consists of a flat plane surface represented using a triangle mesh. All three experiments, i.e., visual, haptic, and visual–haptic, are conducted using the Reachin Display (see Fig. 1) [24], i.e., a vision–haptic system that integrates a haptic device with stereo graphics for an immersive and high-quality 3-D experience.

The importance of using a colocated visual–haptic setup with stereoscopic information as opposed to a noncollocated display (i.e., the workspace of the haptic device is located outside the visual area) has been discussed in several research studies [25]–[29]. In [26] and [28], the authors emphasized the relevance of vision–haptic collocation in improving a user's perception of depth when interacting with virtual environments. Similarly, in [27] and [29], several experiments were conducted to demonstrate that colocation of the visual and haptic sensory

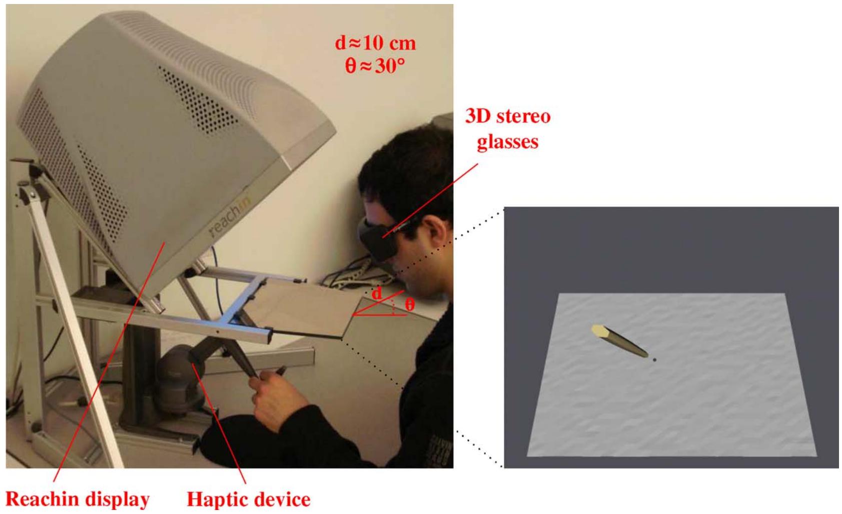


Fig. 1. Participant's position and orientation with respect to the Reachin Display (left), as well as an example of the (watermarked) haptic-enabled virtual environment presented to the user (right) [3].

conditions does in fact lead to improved task performance and enhanced sense of presence within a virtual environment. More importantly, in [27], the authors revealed that collocation is of great benefit to tasks requiring accurate positioning (which is of crucial significance when conducting haptic watermarking experiments). In addition, in [25], Congedo *et al.* observed that for spatial tasks, delocation of visual and haptic percepts prevented a balance in the integration of the two modalities by favoring the dominant sense. For the foregoing reasons, and to obtain reliable and high-precision vision–haptic results, all our experiments were conducted in collocated conditions using the Reachin Display. The Reachin Display's high-accuracy and collocated visual–haptic setup has made it very popular in various precision-demanding applications, specifically in medical training, where surgical simulations of complex procedures are typically required.

The haptic stimulus was sensed using the SensAble PHANTOM Desktop force-feedback device, which is equipped with an encoder stylus that provides six-degree-of-freedom single-contact-point interaction. The PHANTOM Desktop device is reputed for its low backdrive friction (< 0.23 oz), high-fidelity force feedback output, and high-resolution positional sensing (nominal position resolution of 0.02 mm), making it ideal for high-precision haptic experiments. Furthermore, to guarantee a "real" multimodal visual–haptic sensing of watermarked 3-D objects, the Reachin Display (including the haptic device) was periodically recalibrated to ensure a near-perfect collocation between the virtual environment projected on the mirror and the actual physical position of the haptic device (the collocation error is sufficiently small for it not to be perceived even by human subjects with acute senses).

A. Subjects

Altogether, 12 participants (six males and six females, aged 23–28), 11 right handed and one left handed, with no known sensory motor impairments with their hands took part in the experiments. Their prior experience with PHANTOM haptic devices ranged from novice to expert. All participants were remunerated for their time.

B. Haptic Watermark Embedding and Detection

The watermark is embedded in the haptic-enabled virtual surface by perturbing the coordinates of the mesh vertices. This is achieved using the following method:

$$H_w(i) = H(i) + \beta \cdot w(i) \cdot n(i) \qquad (1)$$

where

$$H(i) = [x \quad y \quad z]^T \quad n(i) = [0 \quad 0 \quad 1]^T.$$

More precisely, $H(i)$ corresponds to the ith vertex of the host plane, whereas $H_w(i)$ denotes the ith vertex of the watermarked plane. $w(i)$ is an independent and identically distributed watermark embedded along the surface normal $n(i)$ and consists of random numbers that follow a Gaussian distribution with zero mean and unit variance. Finally, β is a parameter that controls the watermark embedding strength. Conversely, the watermark detection is performed by 1) looking at the plane; 2) touching the plane; and 3) looking at and touching the plane.

C. Conditions

There were three experimental conditions in our analysis that depended on the type of sensory feedback presented to each participant during each experiment. The following sensory conditions were used for our study:

1) only visual feedback;
2) only haptic feedback;
3) visual and haptic feedback.

Under all three conditions, for the sake of consistency, the participants were requested to comfortably sit in front of the Reachin Display at a distance $d \approx 10$ cm and an orientation $\theta \approx 30^\circ$ from the device's half-mirror (used to align the visual and haptic stimuli and prevent a participant from seeing his or her hand when manipulating the virtual object), as depicted in

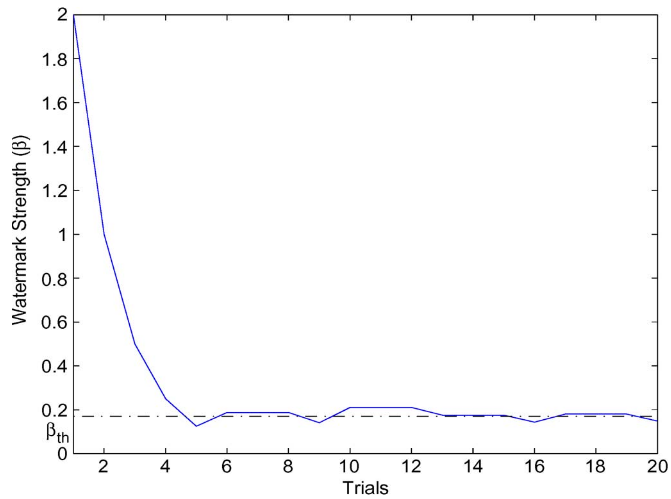


Fig. 2. Typical experimental staircase procedure for one run [3]. The dashed line represents the resulting detection threshold ($\beta_{\rm th}$).

Fig. 1 (left). An example of the virtual environment presented to the user during a visual–haptic experiment is shown in Fig. 1 (right). It encompasses a (clearly watermarked) virtual plane and the virtual stylus used to interact with the surface mesh. Moreover, the color of the surface used in the experiments is in fact blue. A light-gray color is used in Fig. 1 (right) for better quality black and white printouts of the paper. Furthermore, the experiments were conducted in a dark room to avoid any effects of ambient lighting.

D. Procedure

The watermark detection thresholds for each subject, and for the three experiments, were determined using the two-interval forced choice paradigm along with the adaptive staircase (AS) method [30]. In each trial, subjects viewed two stimuli in subsequent intervals and in random order. One of the stimuli was the host plane; the other was the watermarked plane. The subjects' task was to report which of the two stimuli contained the watermark. Trial-by-trial correct-answer feedback was never provided during the experiment. To determine the detection threshold $\beta_{\rm th}$, a "lead-in" one-up one-down AS rule was initially used to speed up the approach to the β region of interest. During this phase, β was halved after each correct response, until the first error occurred. The staircase then followed a one-up three-down rule (β is only decreased after three consecutive correct answers and raised again following a single wrong answer). In this second phase, β was increased or decreased by multiplying it by a factor $k = 1.5$ or 0.75, respectively. After three reversals are obtained, a third phase is initiated (also follows the one-up three-down AS method), where β is increased or decreased by multiplying it by a factor $k = 1.25$ or 0.825, respectively. Trials continued until a total of three reversals were obtained in the third phase of the staircase. Reversals being defined as when the watermark strength changes from increasing to decreasing or vice versa. The detection threshold $\beta_{\rm th}$ is computed while taking the average over the last phase of reversals. As an example, a plot of one of our experiments using the AS procedure is presented in Fig. 2.

V. Results

In this section, detection thresholds resulting from each of the three experiments (i.e., visual, haptic, and visual–haptic perceptibility of the watermark) for every participant, and repeated for different resolutions of the virtual surface, are presented. General observations of the provided results are initially discussed, followed with a detailed statistical analysis of the data.

A. Detection Thresholds

For each of the three experiments, the participants were presented with a lower and a higher resolution version of the virtual surface where the former is represented using 672 triangles, whereas the latter is constituted of 2048 triangles. The plane mesh displayed in all of the undertaken experiments is a rectangular surface of size 24 cm $\times$ 18 cm. Furthermore, haptic rendering (i.e., the synthetically generated haptic stimuli) is performed via the common virtual proxy method [31]. Effort was made to make sure that the subjects are well rested prior to the experiments, as lack of rest could impact the subject's sense of vision and touch, as well as their ability to focus. In addition, participants took breaks as needed during each experiment.

Experiment I—Visual Perceptibility: In this experiment, the participants were requested to detect the watermarked plane while solely relying on visual feedback. Three runs of the experiment were conducted for each of the two available resolutions of the surface, and the corresponding detection thresholds were obtained using the aforementioned AS technique. A box plot of the resulting vision-based detection thresholds of all 12 participants (three detection thresholds are acquired per participant) for the lower and higher resolution meshes are illustrated in Figs. 3(a) and 4(a), respectively.

Experiment II—Haptic Perceptibility: In the second experiment, the participants were requested to detect the watermarked plane while solely relying on haptic feedback. Three runs of the experiment were carried out for each of the two available resolutions of the surface, and the corresponding detection thresholds were obtained using the AS technique. A box plot of the resulting haptic-based detection thresholds of all participants, for the lower and higher resolution meshes, is illustrated in Figs. 3(b) and 4(b), respectively.

Experiment III—Visual and Haptic Perceptibility: As for the visual–haptic case, the participants were requested to detect the watermarked plane while relying on both visual and haptic feedback. Similar to the first two experiments, three runs of the experiment were performed for each of the two available resolutions of the surface, and the corresponding detection thresholds were obtained using the AS method. A box plot of the resulting bimodal vision–haptic-based detection thresholds of all participants for the lower and higher resolution meshes is illustrated in Figs. 3(c) and 4(c), respectively.

B. Initial Observations

To better visualize and analyze the results considering that the obtained detection thresholds across the different modalities are fairly similar, the threshold values of all the

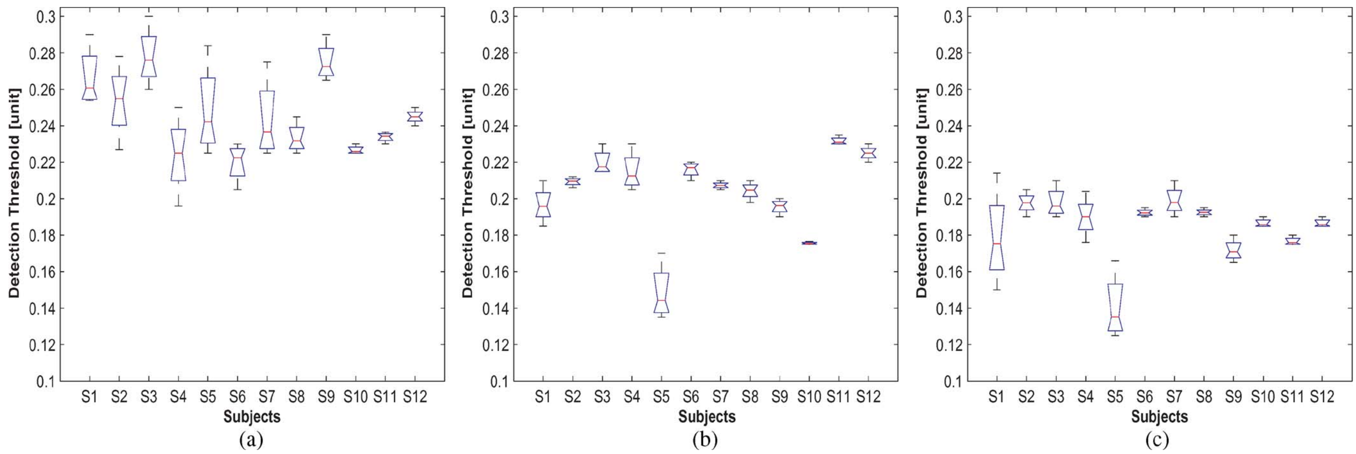


Fig. 3. Results obtained in experiments I, II, and III, where (a) visual feedback, (b) haptic feedback, and (c) visual and haptic feedback were provided. The surface used in these cases is a low-resolution rectangular mesh of size 24 cm × 18 cm represented using 672 triangles.

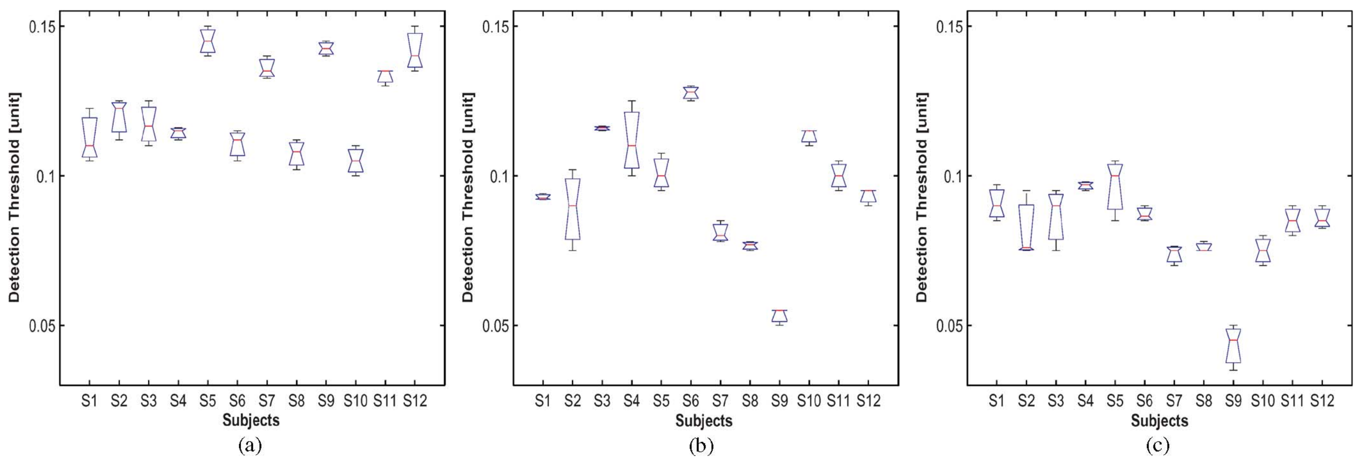


Fig. 4. Results obtained in experiments I, II, and III, where (a) visual feedback, (b) haptic feedback, and (c) visual and haptic feedback were provided. The surface used in these cases is a high-resolution rectangular mesh of size 24 cm × 18 cm represented using 2048 triangles.

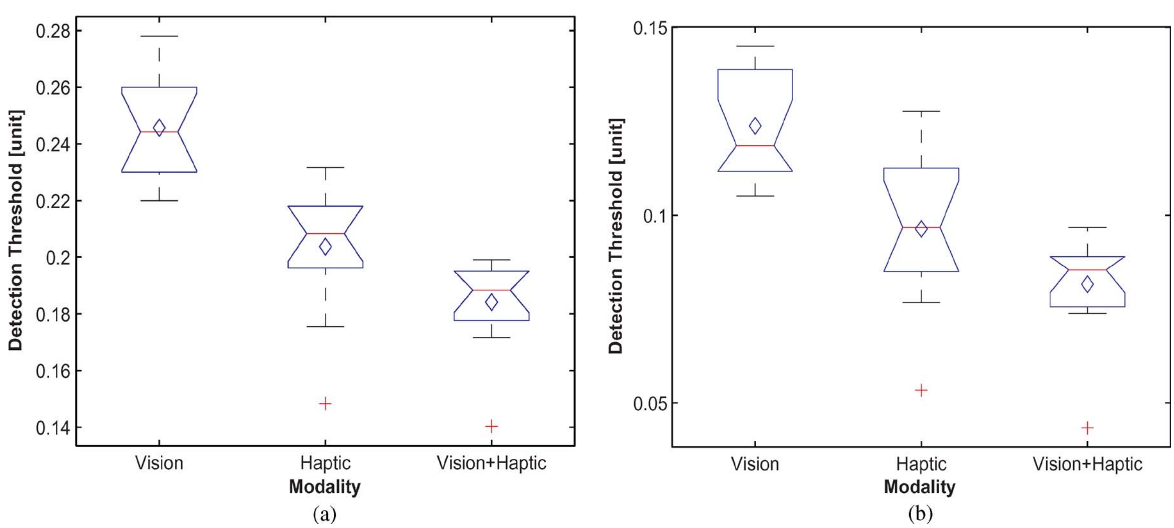


Fig. 5. Average detection threshold values of all the participants for each run of every experiment and the corresponding standard deviations. The results for the low- and high-resolution meshes are presented in (a) and (b), respectively.

participants for each run of every experiment were averaged and plotted for the lower and higher resolution meshes, as shown in Fig. 5(a) and (b), respectively. Several observations can be made from the results presented in Figs. 3–5. First, for all 12 subjects, the average detection thresholds obtained in the haptic-alone experiment are for the most part below the detection thresholds determined when the participants solely relied on visual feedback. This suggests that the watermark is possibly more easily detected when touching a surface mesh as opposed to performing a visual assessment. This assumption is in agreement with the results already obtained in [22], where the perceptibility of a watermark embedded in a 3-D mesh is also inspected using vision-alone and haptic-alone (bimodal visual–haptic experiments were not conducted, and as aforementioned, a statistical analysis was not performed to test for significance between modalities). However, in [22], the results illustrated that the average detection thresholds of the haptic-only experiments are in most cases less than half the threshold values obtained with vision-only. Conversely, in the results presented here, it can be observed from Fig. 5 that the average detection thresholds (mean values are identified with the diamond-shaped marker symbol) of the haptic-only conditions are only approximately 20% less than the threshold values obtained with vision-only. This variation in the results is possibly due to several reasons. First, in [22], a fewer number of participants took part in the experiments, and a smaller number of runs per experiment were performed. These factors can evidently lead to less precise results. Moreover, the haptic-only experiments performed in [22] make use of the same high-precision haptic device used in the experimental setup exploited in this paper. However, their vision-only experiments were performed using a primitive computer display (cathode ray tube monitor). Conversely, the experimental setup presented here relies on stereo graphics for an immersive and high-quality 3-D experience, which in turn enabled us to perform high-precision vision-based experiments. Thus, the large margin between the detection threshold values of the haptic-only and vision-only experiments presented in [22] is probably also due to the considerable difference in quality between their haptic and visual displays.

Another observation that can be deduced from the results illustrated in Fig. 5 is that the detectability of the watermark increases with the resolution of the surface mesh in all three sensory conditions (vision, haptic, and vision + haptic). Moreover, from Fig. 5, it is shown that the between-subject variability (the standard deviations) of the detection thresholds for both experiments and the two presented resolutions is quite high. Conversely, from Figs. 3 and 4, it can be seen that the within-subject variability is relatively low, suggesting that our experimental procedure and setup are accurate. The large between-subject variability, however, is probably due to the fact that when performing a perceptual judgment, participants exploit in a different manner the information available through their distinct sensory channels. In addition, it is known that PHANTOM haptic devices induce low but perceivable backdrive friction, which can negatively affect a user's haptic perception [32]. In turn, during the haptic and bimodal vision–haptic experiments, it is believed that certain observers were not able to isolate the backdrive distortion as well as others when attempting to feel the barely noticeable watermark. This could also have contributed to the relatively large between-subject variability of the results.

Furthermore, in the bimodal case, it seems that vision–haptic feedback is better than any of the single modalities when detecting a watermark embedded in a 3-D mesh. This can best be observed from the results depicted in Fig. 5, where the lowest detection threshold values were obtained in the vision–haptic conditions. This suggests that there is possibly a higher level of discrimination performance resulting from the combined modality condition.

C. Statistical Data Analysis

To further confirm the observations discussed in Section V-B, a two-way within-subject ANOVA is conducted to statistically explore the impact of the considered modalities (vision, haptic, and vision–haptic) and the virtual mesh resolutions (low and high) on the acquired watermark detection thresholds. The two-way ANOVA enables us to simultaneously test for the effect of each of the independent variables (modality, resolution) on the dependent variable (watermark detection threshold) and also identifies any interaction effect. For each participant, the detection thresholds used in the analysis are in fact the average values of the runs carried out for every (modality, resolution) experimental condition, e.g., the detection threshold for Subject 1 (S1) under the experimental condition $[\text{Modality} = \text{Haptic}, \text{Resolution} = \text{Low}]$ is equal to the average of the three threshold values obtained during the three corresponding runs.

Throughout the analysis, when violations of the assumption of sphericity for the repeated measures ANOVA were observed, Greenhouse–Geisser adjustments were used. In addition, pairwise comparisons and Bonferroni adjustments for multiple comparisons were conducted as necessary [33]. For all the experimental conditions, the significance level adopted was $\rho = 0.05$. A statistically significant main effect for *Modality* was observed $F(1.225, 13.480) = 41.4$; $\rho < 0.0005$ with a large effect size (partial eta squared) $\eta_p^2 = 0.79$ [34]. The effect size is specified to indicate the proportion of variance of the acquired detection thresholds that is explained by the main effect (Modality), independent of the number of samples. A posttest pairwise comparison with Bonferroni adjustment indicated a significant difference between the vision and haptic modalities ($\rho = 0.002$), between the vision and vision + haptic modalities ($\rho < 0.0005$), and between haptic and vision + haptic modalities ($\rho < 0.0005$).

In addition, as alleged in Fig. 5, the factor *Resolution* also has a statistically significant main effect, where $F(1, 11) = 445.2$; $\rho < 0.0005$ with a large effect size $\eta_p^2 = 0.976$. However, it was observed that the *Modality* × *Resolution* interaction effect, which is the combined effect of the *Modality* and *Resolution* factors on the detection thresholds, was not statistically significant as $F(2, 22) = 3.097$; $\rho = 0.065$ with a small effect size $\eta_p^2 = 0.220$.

Overall, the results indicate that the lowest watermark detection thresholds were obtained in visual + haptic conditions, as compared with unisensory conditions where only vision or haptic feedback is available. This finding suggests that the addition of a second modality provides information (some maybe redundant and some nonredundant) that improves the detectability of watermarks embedded in 3-D models. Furthermore, if we are to compare the unisensory conditions, then we can state with a fair degree of certainty that haptic-alone conditions dominate vision-alone when detecting a watermark embedded in a 3-D mesh. In addition, although it is very apparent that the detectability of the watermark increases with the resolution of the surface mesh, the low F value for *Modality* $\times$ *Resolution* interaction confirms that the *Resolution* of the surface has absolutely no effect on the exploited *Modality*, i.e., surface resolution and modality conditions are entirely independent. Consequently, the impact of the selected modality on the perceptibility of the 3-D watermark is independent of the chosen surface resolution.

VI. Conclusion

Haptic watermarking is a brand new research challenge that is expected to significantly expand the field of 3-D digital watermarking in the very near future. In this paper, an experimental study has been presented to investigate the role of bimodal vision–haptic feedback in the perception of a watermark embedded in a haptic-enabled 3-D virtual surface. The obtained results suggested that the lowest watermark detection threshold values were obtained when vision and haptic feedback were simultaneously available, as compared with either modality alone. In other words, the participants were able to integrate the input from the two modalities in such a manner to produce a percept that is better than that provided by either single modalities. The presented results are very interesting and encouraging as they clearly demonstrate the importance of the haptic dimension in the perceptibility of a 3-D watermark. Many issues are however left open for further investigation, including the following (some of which stem from the discussion presented in [35]):

1) Investigating the possible impact of different haptic rendering techniques.
2) Different graphic rendering conditions should be analyzed, including surface color, viewpoint, shading, degree of opacity or transparency, lighting, etc.
3) Other more complex shapes of the original 3-D object should be considered.
4) Evaluating the perceptual impact of bimodal auditory–haptic feedback and comparing with visual–haptic conditions. Or of more importance, investigating the role of multisensory vision–haptic–auditory feedback in the perceptibility of a 3-D watermark embedded in haptically enabled meshes.

Finally, this paper is a first, yet a very important, step toward the analysis of multimodal vision–haptic watermarking. Our findings are expected to stimulate the reevaluation of existing mesh watermarking algorithms (using a vision–haptic setup) and will serve as a basis for further studies in haptic digital watermarking.

Acknowledgment

The authors would like to thank the participants for their valuable time and helpful suggestions and the reviewers and the editors for their constructive comments on previous versions of this paper.

References

[1] M. Corsini, M. Barni, F. Bartolini, R. Caldelli, V. Cappelini, and A. Piva, "Towards 3D watermarking technology," in *Proc. IEEE Region 8 EUROCON Conf.*, 2003, vol. 2, pp. 393–396.

[2] D. Prattichizzo, M. Barni, G. Menegaz, A. Formaglio, H. Z. Tan, and S. Choi, "Perceptual issues in haptic digital watermarking," *IEEE Multimedia*, vol. 143, no. 3, pp. 84–91, Jul.–Sep. 2007.

[3] N. Sakr, N. D. Georganas, J. Zhao, and E. M. Petriu, "Perceptibility of digital watermarks embedded in 3D meshes using unimodal and bimodal hapto-visual interaction," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, 2009, pp. 987–991.

[4] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*, 2nd ed. San Mateo, CA: Morgan Kaufmann, 2008.

[5] S. Biswas, S. R. Das, and E. M. Petriu, "An adaptive compressed MPEG-2 video watermarking scheme," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 5, pp. 1853–1861, Oct. 2005.

[6] M. Corsini, E. D. Gelasca, and T. Ebrahimi, "A multi-scale roughness metric for 3D watermarking quality assessment," in *Proc. Workshop Image Anal. Multimedia Interactive Services*, 2005.

[7] P. R. Alface and B. Macq, "Shape quality measurement for 3D watermarking schemes," in *Proc. SPIE-IS&T Electron. Imaging, Security Watermarking Multimedia Contents*, 2006, vol. 6072, pp. 622–634.

[8] M. Paterson, *The Senses of Touch: Haptics, Affects and Technologies*. Oxford, U.K.: Berg, 2007.

[9] X. Shen, J. Zhou, A. Hamam, S. Nourian, N. R. El-Far, F. Malric, and N. D. Georganas, "Haptic-enabled telementoring surgery simulation," *IEEE Multimedia*, vol. 15, no. 1, pp. 64–76, Jan.–Mar. 2008.

[10] A. Alamri, M. Eid, R. Iglesias, S. Shirmohammadi, and A. El Saddik, "Haptic virtual rehabilitation exercises for poststroke diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 9, pp. 1876–1884, Sep. 2008.

[11] M. A. Srinivasan, G. L. Beauregard, and D. L. Brock, "The impact of visual information on haptic perception of stiffness in virtual environments," in *Proc. ASME Dyn. Syst. Control Division*, 1996, pp. 555–559.

[12] W. Wu, C. Basdogan, and M. A. Srinivasan, "Visual, haptic, and bimodal perception of size and stiffness in virtual environments," *ASME Dyn. Syst. Control*, vol. 67, pp. 19–26, 1999.

[13] J. F. Norman, H. F. Norman, A. M. Clayton, J. Lianekhammy, and G. Zielke, "The visual and haptic perception of natural object shape," *Percept. Psychophys.*, vol. 66, no. 2, pp. 342–351, 2004.

[14] J. M. Weisenberger and G. L. Poling, "Multisensory roughness perception of virtual surfaces: Effects of correlated cues," in *Proc. 12th Int. Symp. Haptic Interfaces Virtual Environment Teleoperator Syst.*, 2004, pp. 161–168.

[15] G. L. Poling, J. M. Weisenberger, and T. Kerwin, "The role of multisensory feedback in haptic surface perception," in *Proc. 12th Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2003, pp. 187–194.

[16] M. A. Heller, "Visual and tactual texture perception: Intersensory cooperation," *Percept. Psychophys.*, vol. 31, no. 4, pp. 339–344, Apr. 1982.

[17] S. J. Lederman and S. G. Abbott, "Texture perception: Studies of intersensory organization using a discrepancy paradigm, and visual versus tactual psychophysics," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 7, no. 4, pp. 902–915, Aug. 1981.

[18] S. J. Lederman, G. Thorne, and B. Jones, "Perception of texture by vision and touch: Multidimensionality and intersensory integration," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 12, no. 2, pp. 169–180, May 1986.

[19] K. Drewing, M. O. Ernst, S. J. Lederman, and R. Klatzky, "Roughness and spatial density judgments on visual and haptic textures using virtual reality," in *Proc. 4th EuroHaptics Conf.*, 2004, pp. 203–206.

[20] S. Guest and C. Spence, "What role does multisensory integration play in the visuotactile perception of texture?" *Int. J. Psychophysiol.*, vol. 50, no. 1/2, pp. 63–80, Oct. 2003.

[21] D. Prattichizzo, M. Barni, H. Z. Tan, and S. Choi, "Perceptibility of haptic digital watermarking of virtual textures," in *Proc. World Haptics, 1st Joint Eurohaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2005, pp. 50–55.

[22] A. Formaglio, S. Belloni, G. Menegaz, H. Z. Tan, D. Prattichizzo, and M. Barni, "Perceptibility of digital watermarking in haptically enabled 3D meshes," in *Proc. EuroHaptics Conf.*, 2006, pp. 407–412.

[23] A. Kumazaki, K. Terada, and A. Ito, "Role of vision on haptic length perception," in *Proc. World Haptics, 2nd Joint Eurohaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2007, pp. 336–341.

[24] Reachin Technologies AB. Reachin Display. [Online]. Available: http://www.reachin.se/products/

[25] M. Congedo, A. Lécuyer, and E. Gentaz, "The influence of spatial delocation on perceptual integration of vision and touch," *Presence: Teleoperators Virtual Environ.*, vol. 15, no. 3, pp. 353–357, Jun. 2006.

[26] G. Jansson and M. Ostrom, "The effects of co-location of visual and haptic space on judgments of form," in *Proc. EuroHaptics*, 2004, pp. 516–519.

[27] D. Swapp, V. Pawar, and C. Loscos, "Interaction with co-located haptic feedback in virtual reality," *Virtual Reality J.*, vol. 10, no. 1, pp. 24–30, May 2006.

[28] L. Bouguila, M. Ishii, and M. Sato, "Effect of coupling haptics and stereopsis on depth perception in virtual environments," in *Proc. Workshop Haptic Human-Comput. Interaction*, 2000, pp. 54–62.

[29] S. A Wall, K. Paynter, A. M. Shillito, M. Wright, and S. Scali, "The effect of haptic feedback and stereo graphics in a 3D target acquisition task," in *Proc. EuroHaptics*, 2002, pp. 23–29.

[30] H. Levitt, "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Amer.*, vol. 49, no. 2, pp. 467–477, Feb. 1971.

[31] D. C. Ruspini, K. Kolarov, and O. Khatib, "The haptic display of complex graphical environments," in *Proc. 24th Annu. Conf. Comput. Graph. Interactive Techn.*, 1997, pp. 345–352.

[32] T. M. Massie and J. K. Salisbury, "The PHANTOM haptic interface: A device for probing virtual objects," in *Proc. ASME Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 1994, vol. 1, pp. 295–301.

[33] S. E. Maxwell and H. D. Delaney, *Designing Experiments and Analyzing Data: A Model Comparison Perspective*. Belmont, CA: Wadsworth, 1990.

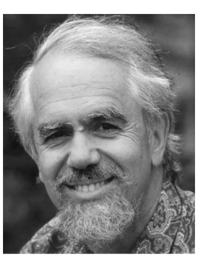
[34] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 5th ed. Boston, MA: Pearson Education, 2007.

[35] S. Belloni, A. Formaglio, G. Menegaz, H. Z. Tan, D. Prattichizzo, and M. Barni, "Is haptic watermarking worth it?" in *Proc. IS&T/SPIE Int. Symp. Electron. Imaging*, 2006, pp. 15–19.

Nizar Sakr (S'06) received the B.A.Sc. degree in computer engineering (*summa cum laude*) and the M.A.Sc. degree in electrical engineering in 2004 and 2006, respectively, from the University of Ottawa, Ottawa, ON, Canada, where he is currently working toward the Ph.D. degree in electrical and computer engineering.

His current research interests include haptics, image processing, and computational intelligence: theory and applications, pattern recognition, and biometrics.

Mr. Sakr is the Chair of the IEEE Computational Intelligence Society, Ottawa Chapter.



Nicolas D. Georganas (F'90) received the Dipl.Ing. degree in electrical engineering from the National Technical University of Athens (NTUA), Athens, Greece, in 1966, the Ph.D. degree in electrical engineering (*summa cum laude*) from the University of Ottawa, Ottawa, ON, Canada, in 1970, the Honorary Dr.-Ing. degree from the Technische Universität Darmstadt, Darmstadt, Germany, in 2004, and the Honorary Dr.-Ing. degree from NTUA, in 2007.

He is a Distinguished University Professor with the University of Ottawa. He holds a Cátedra de Excelencia and is a Visiting Researcher with IMDEA Networks, University Carlos III of Madrid, Madrid, Spain. His research interests are in multimedia communications, ambient computing, and collaborative virtual environments.

Dr. Georganas is a Fellow of the Canadian Academy of Engineering, a Fellow of the Academy of Science (Royal Society of Canada), and a Fellow of the Engineering Institute of Canada. He is a Laureate of the 2002 Killam Prize for Engineering, Canada's highest award for career achievements in research, and an Officer of the Order of Canada. He is the Editor-in-Chief of the *ACM Transactions on Multimedia Computing, Communications and Applications*.

Jiying Zhao (M'00) received the Ph.D. degree in electrical engineering from North China Electric Power University, Beijing, China, and the Ph.D. degree in computer engineering from Keio University, Yokohama, Japan.

He is currently a Professor with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada. His research interests include image and video processing and multimedia communications.

Dr. Zhao is a member of the Institute of Electronics, Information and Communication Engineers and the Professional Engineers Ontario.

Emil M. Petriu (M'86–SM'88–F'01) received the Dipl.Eng. and Dr.Eng. degrees from the Polytechnic Institute of Timisoara, Timisoara, Romania.

Since 1985, he has been with the University of Ottawa, Ottawa, ON, Canada, where he is currently a Professor and the University Research Chair of the School of Information Technology and Engineering. His research interests include robot sensing and perception, interactive virtual environments, human–computer symbiosis, soft computing, and digital integrated circuit testing. During his career, he has published more than 300 technical papers, authored two books, edited other two books, and received two patents.

Dr. Petriu is a Fellow of the Canadian Academy of Engineering and the Engineering Institute of Canada. He is the corecipient of the 2003 IEEE's Donald G. Fink Prize Paper Award and recipient of the 2003 IEEE Instrumentation and Measurement Society Award.