Multimed Tools Appl
DOI 10.1007/s11042-016-3281-x

An information-theoretic treatment of passive haptic media

Konstantinos Moustakas[1] · Aris S. Lalos[1]

Received: 19 February 2015 / Revised: 24 November 2015 / Accepted: 17 January 2016

Abstract Haptic rendering has been long considered as the process of estimating the force that stems from the interaction of a user and an object. Even if this approach follows the principles of natural haptic interaction, it places severe limitations in processing haptic media. This paper presents an information theoretic framework that aims to provide a new view of haptic rendering that can accommodate for open-loop synthetic haptic media, where interaction-based rendering is a special case. As a result, using the proposed information-theoretic approach, the haptic signal can be precomputed as a force field, stored and then filtered by taking into account device and perceptual capabilities of the receiver in order to lower the required bandwidth of the resulting stream, thus opening new possibilities for the representation and processing of haptic media.

Keywords Haptic rendering · Information theory · Haptic information loss · Haptic filter · Haptic coding

1 Introduction

The historical evolution of media clearly demonstrates an increment in the dimensionality of the underlying media sources. Starting from two-dimensional still images, moving to the three-dimensional moving picture, further adding one and two dimensions of mono- and stereo-sound respectively and recently also including the sense of depth in stereoscopic visual representations, currently available and widespread media can be considered as six

✉ Konstantinos Moustakas
moustakas@upatras.gr

Aris S. Lalos
aris.lalos@ece.upatras.gr

1 Electrical and Computer Engineering Department, University of Patras, GR-26500, Rion-Patras, Greece

dimensional. Even if more dimensions can be assumed for different features like color, etc. they can be considered as features of the 6D space.

New dimensions will be inevitably added in the media of the future, possibly including digital representations of signals triggering the human senses, that have not been addressed so far in a systematic way, like touch, olfaction and taste. While all three senses could provide a significant added value in the realism and quality of the future media, research in the haptics domain has progressed significantly to allow for an attempt of systematic and formal definition, representation, processing and rendering of the underlying haptic media [26].

Figure 1, illustrates the complex issue of media source and interrelation, from an information source perspective. On the left part the physical environment is the source of all "direct" media. In particular, light interactions, pressure oscillations and collisions/interaction are considered as the source of "direct visual", "direct audio" and "direct haptic" information respectively. It is evident that all direct media are correlated, since they

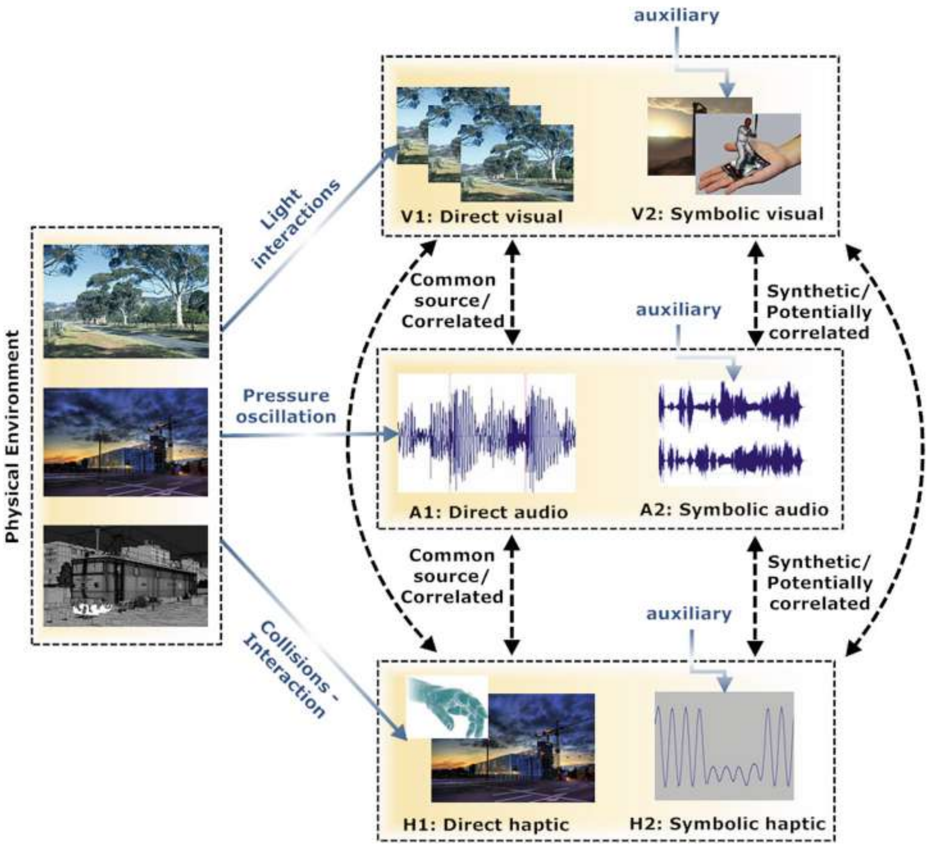


Fig. 1 View of haptics as a medium. *Left*: Physical environment is the source of all "direct" media. Light interactions, pressure oscillations and collisions/interaction are considered as the source of "direct visual", "direct audio" and "direct haptic" information respectively. Auxiliary symbolic media can be added for all communication channels. "Symbolic visual" information includes some forms of visual effects, "symbolic audio" is used to describe audio effects or music, while term "symbolic haptic" information is used to describe the synthetic haptic information, including haptic icons

are due to the same source even if they are caused by different interactions. Moreover, the "media space" is augmented by auxiliary information that is characterized as "symbolic". For example "symbolic visual" information includes some forms of visual effects, overlays or even augmented reality renderings; "symbolic audio" is used to describe audio effects or music, while term "symbolic haptic" information is used to describe the synthetic haptic modelling, including haptic icons [31]. It should be emphasized that even if the symbolic media are not correlated per se, they could exhibit some correlation depending on the particular case.

In general, with the exception of some approaches related to haptic rendering of distance or force fields [21], one of the biggest limitations of current schemes is that haptic rendering is considered only as a result of the interaction of a human user with an underlying (usually 3D) environment. Even if this approach is inspired by the real world, where the sense of touch is triggered by collisions of the human body with the environment, it has three significant drawbacks: i) haptic media cannot be defined as a signal related to a specific physical environment, ii) Off-line processing of haptic media for optimization, compression, indexing, etc. becomes impossible, iii) The requirement for 1kHz update rate, that is considered as the most significant constraining factor imposed on haptic rendering algorithms [15], can be hardly satisfied due to the need for on-line processing of the entire haptic rendering pipeline as also mentioned in a high-level theoretical attempt [2] on haptic information broadcasting media using the existing haptic rendering pipeline.

The proposed framework aims to provide a new view on haptic rendering from an information theoretic perspective. Assuming that the haptic signal can be precomputed as a force field, it can be stored and processed for various processes, including but not limited to coding, indexing, optimization, thus opening new possibilities for the application of haptic media. Moreover, it aspires to provide the necessary theoretical foundations to lead to a compact and holistic representation of haptic media. As a result all three aforementioned limitations will be theoretically overcome for the case of passive haptic media, thus providing a new potential for the field of open-loop haptic rendering. Finally, the proposed approach is instantiated in the context of challenging application scenarios. In particular device and perceptual filters are presented that lead to a first formulation of a haptic media coder, while indicative quantitative examples in the area of haptic media coding are also given.

2 Related work

A very descriptive analysis on the importance of the sense of touch is given in [25] through the effects of its loss that include, among others, inability to eat and walk. While, haptic applications have recently received increasing interest with applications ranging from accessibility engineering [21, 33], entertainment [5, 9] to security and biometrics [3], on contrary to the audio and vision channels, haptic media have not yet been addressed systematically from an information theoretic perspective, leading to non-holistic, fragmented and problem-targeted research.

Haptics research can be divided into three main categories [16]: Machine Haptics, Human Haptics and Computer Haptics [30]. Machine Haptics is related to the design of haptic devices and interfaces, while Human Haptics is devoted to the study of the human perceptual abilities related to the sense of touch. Computer Haptics, or alternatively haptic rendering, studies the artificial generation and rendering of haptic stimuli for the human user. It should be mentioned that the proposed framework takes into account recent research on human haptics [6], while it provides mathematical tools targeting mainly the area of computer haptics.

The simplest haptic rendering approaches focus on the interaction with the virtual environment using a single point [20]. Many approaches have been proposed so far both for polygonal, non-polygonal models, or even for the artificial generation of surface effects like stiffness, texture or friction [12, 15, 20–22]. The assumption, however, of a single interaction point limits the realism of haptic interaction since it is contradictory to the rendering of more complex effects like torque. On contrary, multipoint, or object based haptic rendering approaches use a particular virtual object to interact with the environment. Therefore, besides the position of the object, its orientation becomes critical for the rendering of torques [15, 22].

From an information theoretic perspective, it is worth mentioning that surprisingly coding and compression of haptic data has not been researched extensively so far. Most of the approaches deal with aspects of haptic data transmission in the context of telepresence and teleaction systems [28], focusing mainly on stability and latency issues [8, 13, 24]. Teleoperation of a 6-DoF serial robot using a prorotype haptic device has been proposed in [32]. Differential and entropy coding has been successfully applied in [13], while other traditional approaches including, DPCM ADPCM, Huffman coding have been applied to haptic signals in [13] and [23] respectively.

In [1] predictive coding has been used for haptic signals in order to optimize the communication in teleaction systems with respect to sampling rate, while in [7, 14] perceptual coding mechanisms are developed for haptic data streams. Similarly, human perception-based data reduction schemes for haptic communication in Six-DoF telepresence systems have been proposed in [27, 29], whose properties have been also validated in user studies. Moreover, an interesting approach for modelling haptic vibration textures, from a signal processing perspective, using IIR filters, that can simulate both stochastic and patterned textures of objects, has been proposed in [4].

Finally, an approach to model haptic media as an extension to currently available audio-visual standards is presented in [2], while a first attempt to model haptic rendering from an information theoretic perspective has been presented in [18]. The major limitation of the current approaches is that, since they are targeting mainly telepresence and teleaction systems, they refer to representation and coding of single haptic timeseries that correspond to the force that should exert a remotely manipulated device.

3 Haptic rendering from a multimedia perspective

The potential use of entropy, joint entropy, conditional entropy and mutual information for the case of haptic media has been previously described in [18]. In the following, derivations from this formulation are discussed.

3.1 Information theoretic quantities

It is reasonable to assume that in most applications the reference space or virtual space is of different size compared to the haptic workspace. Let us now assume a mapping M of the haptic information of the reference space into haptic information of the haptic workspace. The latter maybe restricted by several parameters including haptic display limitation on the workspace size, resolution, force exertion amplitude limitations, arithmetics, etc. Let us now define as Workspace Capacity $C(M)$ the average amount of information that the specific haptic workspace setting can render. Since the above parameters are usually constant $C(M)$ is the entropy of a random variable of the specific mapping M. For a specific input haptic space W, following the principle of the data processing inequality the following equation holds:

$$C\left(M\right) = \min\left(C\left(M\left(W\right)\right), H\left(W\right)\right) \tag{1}$$

where the min function encodes the fact the mapping is passive, i.e. no energy/information can be generated through the mapping procedure. This is necessary to avoid active mappings like in rare cases, where the haptic reference space is less detailed than the haptic workspace itself.

Since the quantity C has the same unit as entropy H, we can define the following measures:

$$Workspace\ Mapping\ Ratio\ (WMR) = \frac{C\left(M\right)}{H\left(W\right)} \tag{2}$$

$$Haptic\ Inf.\ Loss\ (HIL) = \frac{\max\left(H\left(W\right) - C\left(M\right), 0\right)}{H\left(W\right)} \tag{3}$$

The above measures are meaningful under the obvious constraints $H\left(W\right) > 0$ and $C\left(M\right) > 0$. The WMR is an index of the amount of information that is transferred for haptic rendering through the haptic workspace, while the HIL encodes the amount of information that is lost in the above procedure.

3.2 Haptic filters

Let us know consider a two dimensional haptic signal -a haptic image- $x(n_1, n_2)$. Without loss of generality assume a typical finite impulse response (FIR) filter $h(n_1, n_2)$. The filtered version of the haptic signal is obtained through convolution as follows:

$$y\left(n_1, n_2\right) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} h\left(k_1, k_2\right) x\left(n_1 - k_1, n_2 - k_2\right) \tag{4}$$

where N_1 and N_2 correspond to the length of the FIR filter.

Now the frequency response of the haptic FIR filter can be trivially obtained through:

$$H\left(\omega_1, \omega_2\right) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} h\left(n_1, n_2\right) e^{-i\left(\omega_1 n_1 + \omega_2 n_2\right)} \tag{5}$$

Let us now discuss a potential application of a haptic filter following the formulation of stand-alone haptic media proposed in this paper. Assuming, without loss of generality, a 2D static signal, i.e. a haptic image, that assigns the magnitude of the force value (scalar force) to each haptic pixel, we can follow the notation of (4). Considering also a tactile T-shirt embedded with an array of $N \times N$ vibrators that encode force magnitude with vibration amplitude. The vertical and horizontal distance of the vibrators is d. Thus, the maximum frequency that can be "rendered" is $f = 2\frac{1}{d}$. Therefore, even if the haptic signal $x(n_1, n_2)$ has a much larger bandwidth, it cannot be displayed using the specific tactile cloth. As a result, using a low-pass filter $H\left(\omega_f, \omega_f\right)$ with cut-off frequency f and convolving its impulse response h_f with x would result in a filtered version of the haptic signal $y_f = x * h_f$ with much lower entropy, i.e. less information content. In other words, following the above procedure the redundant information that cannot be displayed is filtered out and the haptic signal can be encoded much more efficiently in a prospective haptic media transmission system.

Following a similar methodology different implementations of popular filters can be used for the processing of spatio-temporal haptic signals, like optimal quincunx sampling and filtering for the generations of multiresolutional representations of haptic signals.

Additionally, several haptic properties can be considered as haptic filters [20]. Gaussian kernels, or even anisotropic diffusion for smoothing can result in an information theoretic way to perform off-line force shading [20]. Moreover, a salt-and pepper haptic texture can be synthesized by convolving the force field with a high-pass kernel. Friction can be simulated by transmitting only the friction coefficient as metadata with the haptic stream and estimating the normal vector of the surface directly from the force field [21]. Damping can be also similarly simulated using the damping coefficient and the velocity of the haptic interaction point.

4 Instantiation: compressing haptic media

The proposed haptic media representation is general, holistic and can accommodate interaction-based rendering as a special case. The potential applications of the framework are numerous, including typical multimedia operations like off-line processing, haptic editing, optimization, indexing, similarity estimation and Level-of-Detail (LoD) representation.

Additionally, the aforementioned theoretical quantities can lead to optimal Level-of-Detail (LoD) haptic rendering and compression, similarity estimation and design of device and perceptual haptic filters necessary for haptic media coding. Other interesting applications include optimal adaptive workspace partitioning in disjoint subspaces so as to minimize information loss taking into account the rendering device or even correlation analysis between haptic media and visual media, etc. In the following, examples on formulating haptic media compression problems using the proposed framework are provided.

Haptic media are assumed to be defined within the input haptic space W that refers to a virtual environment augmented with haptic information. W actually defines a mapping of the $\Re^3$ Euclidean space into the $\Re^3$ force field that refers to the force exerted to a point object lying on the specific point in space. It should be emphasized that in the general case the haptic space evolves over time and therefore the static W representation becomes a timeseries W_t.

4.1 Spatial redundancy

It is apparent that since haptic media are defined in $\Re^3$ compared to $\Re^2$of the visual media, their storage complexity becomes in the general case $O(n^3)$. However, unfortunately haptic displays have not evolved similarly to visual displays and therefore the spatial resolution they can provide to the user is extremely low with respect to the sensing potential of humans. This is reflected to a very low workspace mapping ratio (2) that can be rendered using typical devices. In information theoretic terms we call this effect a device filter that limits the amount of information that can be perceived as introduced and described in [18].

4.2 Temporal redundancy and perceptual haptic filters

Besides haptic device display capacity, the perceptual capabilities of the human with respect to temporal resolution have specific limitations [7, 8, 14] that can be exploited in order to compress haptic media.

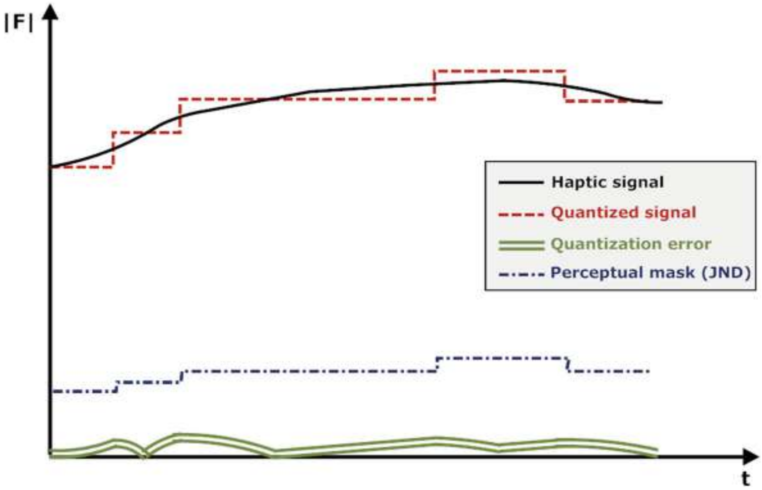


Fig. 2 Qualitative example of a haptic signal $\mathbf{x}$, its quantized version $Q(\mathbf{x})$, the quantization error $|Q(\mathbf{x}) - \mathbf{x}|$ and the JND threshold (Perceptual mask) $\Delta(\mathbf{x}) = k \cdot \mathbf{x}$ for $k = 0.25$. x-axis represents time $\mathbf{t}$, while y-axis force magnitude $|\mathbf{F}|$. The adaptive quantization step has to be carefully selected so that the quantization error never exceeds the perceptual mask, thus leading too perceptually lossless encoding

Some approaches presented in the literature, focusing on telepresence and teleaction [7], tackle the problem of perceptual coding. For the specific case of point-based interaction, they aim to define masking thresholds of force differences between consecutive frames beyond which a difference in the force fed back to the user cannot be perceived.

The complexity of this problem for the general case of haptic media and general multi DoF haptic devices is very high. It does not only depend on the temporal relation between two successive force stimuli but also on the body part that they are applied and the spatial proximity of concurrent stimuli. Since, a detailed analysis is out of the scope of present paper, without loss of generality, let us consider the "deadband" approach described in [7].

In [7] a perceptual mechanism for haptic data compression is proposed based on Weber's Law and the "Just Noticeable Differences" (JND) principle. In particular, for a specific time instance, the respective haptic sample is transmitted only if it can be perceived by the user, i.e. if it lies outside the perceptual mask defined by the previously transmitted samples.

This procedure can be seen as a *perceptual filter* that is applied on the input data stream modifying its entropy. Then all quantities described in Section 3 can be applied on the perceptually filtered input. Now the design problem lies on the definition of a perceptual, multivariate in the general case, filter that takes into account the potential haptic rendering schemes, in terms of available degrees of freedom, perceptual correlation between them, etc.

Now, consider a continuous haptic signal $\mathbf{x}$. The following process results in a scalar quantization $Q(\mathbf{x})$ of $\mathbf{x}$:

$$Q(\mathbf{x}) = \operatorname{sgn}(\mathbf{x}) \cdot \Delta_{JND} \cdot \left\lfloor \frac{|\mathbf{x}|}{\Delta_{JND}} + \frac{1}{2} \right\rfloor \tag{6}$$

where $sgn()$ is the sign function and Δ_{JND} the "Just Noticeable Differences" threshold. Now taking into account that the JND threshold usually depends on the magnitude of the haptic signal then Δ_{JND} should be a function of $\mathbf{x}$ thus leading to $\Delta_{JND}(\mathbf{x})$. The transformation of (6) outputs a haptic signal that is optimal in terms of minimizing the perceptual temporal information overhead of the haptic signal $\mathbf{x}$.

Figure 2 qualitatively illustrates, in the context of a "time"-vs-"force magnitude" graph a haptic signal $\mathbf{x}$, its quantized version $Q(\mathbf{x})$, the quantization error $|Q(\mathbf{x}) - \mathbf{x}|$ and the JND threshold $\Delta(\mathbf{x}) = k \cdot \mathbf{x}$ for $k = 0.25$ [14]. It should be emphasized that in a potential quantizer, the adaptive quantization step has to be carefully selected so that the quantization error never exceeds the JND threshold as illustrated in Fig. 2.

However, force is not a scalar but rather a 3 dimensional vectorial quantity. A more appropriate quantizer for this purpose is a vector quantizer. Let B_l be the partition regions of the 3D space and $\mathbf{g}_l$ the reconstruction vectors or codewords that correspond to each region B_l. Then the quantization function can be formulated as follows:

$$Q_V(\mathbf{x}) = \mathbf{g}_l, \quad \forall \mathbf{x} \in B_l \tag{7}$$

Now the problem of designing a perceptual filter is reduced to i) the generation of the optimal partitioning of the 3D feature space that actually models all forces that can be potentially applied and ii) to the assignment of a reconstruction vector to each partitioned cell.

A potential optimal solution should make sure that the transition between two "neighbouring" reconstruction vectors should just fall within the perceptual masking threshold in a similar way performed in the scalar quantization case (Fig. 2).

As an example let us consider the concentric tessellated icosahedra illustrated in Fig. 3. Each vertex of the polyhedra corresponds to a potential reconstruction vector $\mathbf{g}_l$. In this case the areas B_l are the Voronoi cells of the Voronoi diagram created by the polyhedra vertices. It is apparent that the radius difference between consecutive polyhedra is related with parameter $\Delta_{JND}(\mathbf{x})$ since for a constant force direction a change in the radius reflects change in the force magnitude. On contrary, the distance between neighbouring vertices of the same polyhedron denoted as $\Delta_V(\mathbf{x})$ is related to the Just Noticeable Differences threshold related to a change in rather the force direction than magnitude. Even if it has been reported [14] that the $\Delta_V(\mathbf{x})$ threshold can be modelled as $\Delta_V(d\mathbf{x}) = k \cdot |d\mathbf{x}|$, where $d\mathbf{x}$ is the difference vector for two consecutive force instances, this remains a challenging open research topic.

Similar schemes can be also derived for the spatial masking case, since it is well known that perception of a haptic signal in a specific body position is perceived differently or even not perceived at all in the presence of other stimuli in its close proximity.

Now the usefulness of such device and perceptual filters or processes is obvious when haptic media coding is concerned, since a filtered haptic signal exhibits much lower entropy, without sacrificing the quality of the perceived haptic information, and can be thus coded much more efficiently.

4.3 Haptic media coding

A direct application of the proposed information theoretic framework for open-loop haptic rendering is coding of streaming haptic media. Taking also into account that haptic coding is one of the major milestones a haptic media framework has to reach, an architectural example of a prospective haptic media coder is presented below.

On contrary to existing approaches that mainly encode and transmit a force time-series [14], the problem of coding stand-alone haptic media as defined herein is much more complex. The components of a haptic media coding algorithm are defined to a large extent by

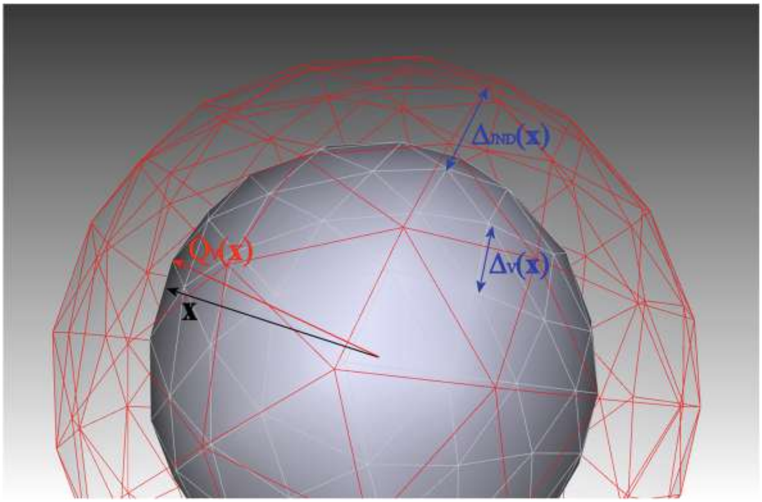


Fig. 3 Concentric tessellated icosahedra for vector quantization. Each point of the polyhedra corresponds to a potential reconstruction vector $\mathbf{g}_l$. The radius difference between consecutive polyhedra is related with parameter $\Delta_{JND}(\mathbf{x})$, while the distance between neighbouring vertices of the same polyhedron denoted as $\Delta_V(\mathbf{x})$ is related to the JND threshold that refers to a change in rather the force direction than magnitude

the source model that is adopted for modeling the haptic media that may make assumptions about the spatial and temporal correlation between neighbouring samples of a haptic media sequence. Figure 4 illustrates a basic architecture of a haptic media coder.

In the encoder the haptic media stream is initially described using the parameters of the source model. For example if a source model of statistically independent force samples is adopted then the parameters of the source model would be the components of the force vector. On the other hand, if a model is adopted that describes a scene as several objects, then the parameters will be force vectors that correspond to each specific object. In the next step, the redundant information is reduced, using device and perceptual filtering as described in Section 4. It should be emphasized that this step results in theoretically lossy, but perceptually loss-less compression. In the next step the filtered parameters are quantized to into a finite set of symbols according to the desired trade-off between bit-rate and distortion. Finally the quantized parameters are finally mapped into binary codewords using coding

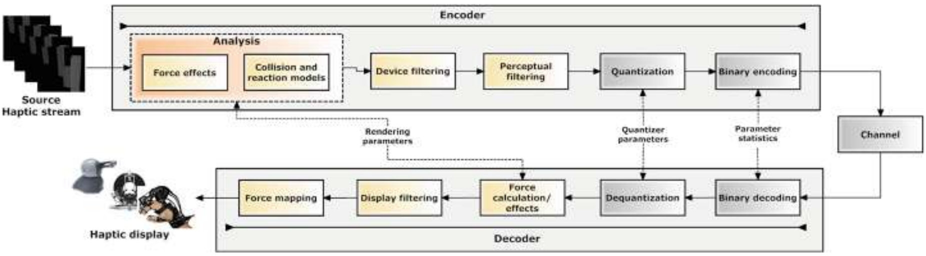


Fig. 4 A basic haptic media coding pipeline. For the “quantization”, “binary coding” as well as their inverse modules general purpose SoA approaches can be used, even if taking into account several properties of the haptic media signals could improve rate-distortion performance. The rest of the modules refer specifically to haptic media

techniques that exploit the statistics of the quantized parameters. The resulted stream is transmitted over the communication channel.

The decoder, by reversing the binary encoding and quantization procedures, retrieves the quantized parameters of the source model and then proceeds to the synthesis of the haptic stream that is subsequently mapped and rendered on a specific haptic interface. The quantization and binary encoding/decoding blocks are typical to any coding scheme, while the rest refer strictly to haptic media. Moreover, source model parameter quantization can be easily omitted since the perceptual filtering already performs quantization below the masking threshold as described in Section 4.

5 Application examples

A direct comparison of the proposed scheme with typical single-contact point, closed-loop system typically used in tele-operation is not fair since the proposed scheme is a) multi-contact, b) open-loop, and totally different from a) single-contact point, b) closed-loop systems. Each one of these approaches is clearly superior in their own field. I.e. when it comes to applications using single-contact point haptic devices, where interaction with the real/virtual environment is critical (e.g. tele-medicine, robotic manipulation) then the closed-loop scheme should be adopted. On contrary, in cases where the haptic medium is there and has to be delivered passively to the user (e.g. haptics-enabled television, real-time haptic sensing and navigation for the blind), then the open-loop information theoretic framework can be clearly considered more appropriate.

Moreover, a detailed experimental evaluation of the proposed theoretical framework is out of the scope of this paper that mainly aims to introduce a different theoretical viewpoint of haptic rendering. However, in order to provide some indicative concrete examples on how the proposed metrics, as defined in Section 3, could be used, some indicative experiments have been performed and are presented below.

Following the approach described in Section 4.3, a simplified haptic media coding scheme can be easily derived. Referring to the block diagram of Fig. 4, we present both a qualitative analysis that ignores the "quantization", "binary coding" as well as their inverse modules, followed by a quantitative one that considers all the block of the presented architecture. Both studies focus on the effects of device/perceptual filtering and force mapping.

5.1 Qualitative analysis

Let us know consider the "Scene 1" of Fig. 5a that is comprised of four objects of low complexity. The objects are left to fall on the ground in a physics-based manner, thus resulting in a two-second simulation. As a result and in order to assure the 1kHz rate for haptic rendering, 2000 updates have been computed for the simulation. Four of them are depicted in Fig. 5a-d. The resulting force fields are generated using the approach described in [21].

"Scene 2" is based on a 20 frame monoscopic image sequence, where dense depth maps are extracted using the approach of [19]. A frame of the sequence and the corresponding depth map is illustrated in Fig. 6. Moreover, a 800 frame depth map sequence is generated using linear interpolation of the depth maps. The major difference of the second scene compared to the first is that it refers to video+depth data that can be rendered in terms of haptics as a height map. The force field that will be subsequently analyzed is generated using the approach described in [21], i.e. a time varying force-field is generated for every

(a) (b) (c)

Fig. 5 Scene comprised of seven low complexity objects. The objects are left to fall on the ground in physics-based manner. For each frame of the simulation a time-instance of the force field is considered that encodes force magnitude and direction for every point of a dense spatial sampling of the scene. Four instances of the simulation are depicted above

time instance of the generated reconstructed scene.

Now, let us assume two different types of haptic displays; Type 1: a 3DoF point display (e.g. Phantom Omni); Type 2: a tactile t-shirt equipped with a $N \times M$ array of force actuators. The major difference of the above two cases, of interest in the following discussion, is that for case 1 the potential user is free to explore the environment, while in case 2 each actuator is linked to a specific area-point of the scene and renders the respective force magnitude, no matter the position or motion of the user, following a similar paradigm to video broadcasting. Thus the two rendering schemes, implied by the aforementioned device types, exhibit a significant difference related to the possibility to navigate the environment. This interesting issue is further discussed in Section 6.

Consider, now a haptic signal defined in a quantization $W \in \Re^3$. In particular, assume a uniform 3D workspace sampled with 10^4 samples per dimension, thus leading to 10^{12} samples. The haptic signal related to the simulation illustrated in Fig. 5, is defined over the quantization W. It is evident that for each one of the 2000 updates of the haptic thread of the simulation there is a need to store, process, transmit 10^{12} force vectors.

Table 1 presents the representation/storage requirements of the above haptic signal after being processed with some of the tools proposed in previous sections and using a haptic display of Type 1. The columns "voxel" and "object" refer to the representation model adopted. In the "voxel" case, the full haptic scene is represented by the regularly sampled voxels,

(a) (b)

Fig. 6 "Scene 2" (**a**) Image and (**b**) Depth map

Table 1 Effect of device and perceptual filters on the two processed scenes for a device of type 1

	Scene 1		Scene 2	
Type 1 device	voxel	object	voxel	object
Size (1^{st} frame)	10^{12}	$2.2 \cdot 10^7$	10^{12}	$6.5 \cdot 10^{10}$
Device filter (1^{st} frame)	$8 \cdot 10^9$	$1.8 \cdot 10^5$	$8 \cdot 10^9$	$5.5 \cdot 10^8$
Entropy (1^{st} frame)	1.2	2.3	3.7	5.2
Raw storage per sample	96 bits	96 bits	96 bits	96 bits
Perceptual filter	18 bits	18 bits	18 bits	18 bits

while the "object" representation follows the typical MPEG-4 object based representation format, meaning that a haptic scene is comprised by its haptic objects and empty space. If in a particular place there is no haptic object then no force should be rendered. It is clear that the "object" representation is more compact per se, even if it requires a clear segmentation of the scene that is in general an ill-posed problem.

The first line of the table "Size" reports the size of the raw signal; the second "device filter" the size resulted after applying the device filter of Phantom Omni as described in Section 4; the third line reports the entropy of the haptic signal. The two last lines of the table depict the compression power of a very simple perceptually loss-less perceptual filter as defined in Section 4. Line "Raw storage per sample" describes the amount of bits necessary to describe a sample prior to perceptual filtering, while line "Perceptual filter" reports the amount of bits per sample necessary after perceptual filtering. If statistical binary coding is applied on the haptic media timeseries then the values of the last two lines can be further decreased. Moreover, the value of the entropy in the object representation is higher than in the voxel-representation since only "object samples" are considered, while the empty space is not processed at all. In other words the entropy is calculated based on fewer samples and is therefore artificially escalated.

Similarly Table 2 summarizes the aforementioned values as far as a haptic display of Type 2 is concerned. It is evident that using a device of Type 2 the respective device filter reduces significantly the amount of information to be transmitted, since only a set of $N \times M$ haptic timeseries are of interest and can be potentially perceived.

A very interesting observation is that the entropy of the second frame till the last one is on average 1.6 and 0.8 as far as object-based and voxel-based representation is concerned

Table 2 Effect of device and perceptual filters on the two processed scenes for a device of type 2

	Scene 1		Scene 2	
Type 2 device (32×32)	voxel	object	voxel	object
Size (1^{st} frame)	10^{12}	$2.2 \cdot 10^7$	10^{12}	$6.5 \cdot 10^{10}$
Device filter (1^{st} frame)	1024	64	1024	672
Entropy (1^{st} frame)	0.57	4.6	1.9	3.1
Raw storage per sample	32 bits	32 bits	32 bits	32 bits
Perceptual filter	6 bits	6 bits	6 bits	6 bits

respectively. This is caused by the fact that in subsequent frames there are many samples that need zero bits for their representation since their respective force value falls within the perceptual mask. Additionally, if predictive coding is applied then the entropy of the force error values becomes even lower.

The analysis above provides simple examples on how the proposed framework can be used for haptic media compression of a synthetic and a real image sequence. It should be emphasized that after applying the proposed filters, typical state-of-the-art techniques, like predictive coding, spectral coding, sparsity coding (e.g. Run-Length-Encoding) could be implemented so as to further compress the resulting bit-stream. These processes refer to the grey blocks "quantization", "binary encoding" of Fig. 4.

Moreover, in order to give one more dimension of the application potential of the proposed scheme, let us reconsider a 2D video and two currently existing video processing steps that result firstly to extraction of saliency based features of the videos that are mapped to tactors as described in [11] - this case closely resembles the tactile t-shirt paradigm mentioned before - and automated generation of the force field timeseries as described in [21]. These two cases are totally inline with the Fig. 1, where case 1 corresponds to symbolic haptic media, while case 2 to direct haptic media.

It is evident that while case 2 is computationally intensive and should be performed off-line thus leading to media that should be transmitted, case 1 could be also performed on-line. However, even case 1 generates a tactile field that is mapped on a specific device. Thus, if the approach aims to be device independent a specific tactile field time-series has to be generated. The potential use of the proposed framework for case 2 is clear. Moreover, even for case 1, where visual information is transformed into synthetic tactile cues, the proposed scheme could provide the necessary mathematical framework to perform processes like, rendering of haptic properties, haptic data indexing, filtering, search and retrieval, since the haptic media are treated in an information theoretic manner.

5.2 Quantitative analysis

The focus of this study is to complement the qualitative analysis by evaluating the presented passive haptic media coding schemes in terms of both compression efficiency and reconstruction quality. In this section we consider both the "quantization", "binary coding" as well as their inverse modules, that were not taken into account in the qualitative

(a) (b)

Fig. 7 (**a**) 'Barcelona hotel' scene taken from the SUN3D database (**b**) Corresponding depth map analysis. The effects of the device filter, the perceptual filter and the quantization resolution are extensively studied by using depth images from the SUN3D dataset consisting of full 3D spaces, scanned with an RGB-D sensor [10]. We consider a schene that is based on a frame depth map sequence from the ”hotel barcelona” scans. A frame of the sequence and the corresponding RGB image are illustrated in Fig. 7.

In particular, we assume a uniform 3D workspace sampled with 10^3 samples per dimension, thus leading to 10^9 samples. To simulate the effects of the device filter we assume that the nominal position resolution and the force feedback workspace of the haptic display result into 2^{n_w} samples per dimension. The parameter n_w determines the spatial resolution of the haptic device and as a result, the workspace mapping ratio that can be rendered, which can be evaluated directly from (2).

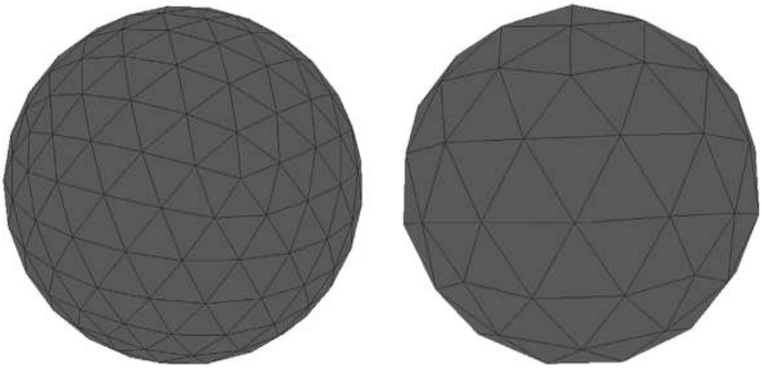
The force fields that will be subsequently analyzed are generated using the approach described in [21], for every time instance of the generated 3D workspace. In order to reduce the transmission requirements we consider only the non zero force vectors. The amplitude and the direction of each non zero force field vector are formulating a vector and a Matrix denoted by **a** and **D** of sizes L and $L \times 3$ which are quantized individually. The direction of each force vector $D_{i,:} \in \Re^{1\times 3}$, $i = 1, \ldots, L$ is quantized using the approach described in Section 4.2. More specifically, let B_l be the partition regions defined by a concentric tessellated icosahedron, where each point $i = 1, \ldots, N_d$ of the polyhedron corresponds to a potential reconstruction vector g_i that is represented by index i. We consider two different polyhedra with $N_d = 258$ and $N_d = 92$ that provide different direction resolutions and are shown in Fig. 8. Then the quantization is performed by using the maximum likelihood approach and mapping an index i_d to each direction $\mathbf{d}_i$ that corresponds to vector $\mathbf{g}_{i_d}$.

Regarding the amplitude, we consider the deadband approach described in [7], which uses a perceptual mechanism for haptic data compression based on Weber’s Law. More specifically, for a specific time instance, the respective haptic sample is transmitted only if it can be perceived by the user, i.e. if it lies outside the perceptual mask defined by the previously transmitted samples. If a_i denotes the amplitude that will be transmitted in time instant i then:

$$a_i = \begin{cases} 0 & if\ a_{i-1} - k \times a_{i-1} < a_i < a_{i-1} + k \times a_{i-1} \\ a_i & otherwise \end{cases} \tag{8}$$

where k is the applied dead band parameter. The samples that are selected for transmission are then quantized using function $Q : \Re \rightarrow \mathbf{Y_i}$ which is a scalar quantization function[1] that discretizes its input, by performing a mapping of each real element of $\mathbf{a_i}$ to a finite set of codewords $\mathbf{A_i}$. For the presented study we used the Lloyd max algorithm [17].The non zero force amplitudes and directions corresponding to a single & multi-depth frame are formulating two individual vectors for transmission that are further compressed by applying loss-less Huffman coding. The Huffman encoding approach generates the binary encoding vectors denoted by C_a and C_d respectively.

[1]Typical quantizers are usually optimized by selecting decision boundaries and output levels in order to minimize the distortion (e.g., mean square error) between the input real number and its quantized representation.



(a) (b)

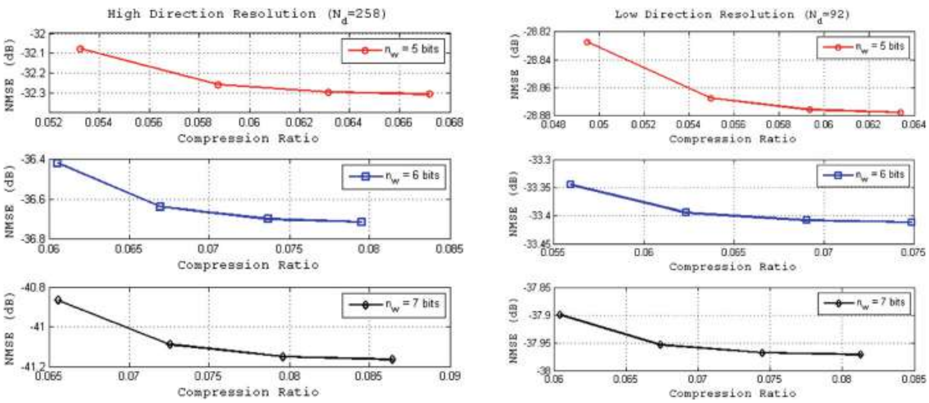
Fig. 8 Concentric tessellated icosahedra for vector quantization of force directions with N_d vertices (**a**) high resolution case ($N_d = 258$ vertices) (**b**) low resolution case ($N_d = 92$ vertices)

The aforementioned approaches (assuming single and multiple frame depths) are evaluated in terms of both compression efficiency and reconstruction accuracy. The compression efficiency is evaluated by using the Compression ratio (CR):

$$CR = \frac{C_a + C_d}{m \times 2^{n_w} \times 2^{n_w} \times 2^{n_w} \times 3 \times 32} \tag{9}$$

where m is the number of processed frames, while the reconstruction effectiveness is evaluated by the normalized mean square error ($NMSE$) defined as:

$$NMSE = \frac{\sum_{j=1}^{L} \left\| a_j D_{j,:} - a_{q_j} g_{j_d} \right\|}{\left\| a_j D_{j,:} \right\|} \tag{10}$$



(a) (b)

Fig. 9 Compression of haptic media generated from a single frame depth. NMSE vs CR for the (**a**) high resolution case ($N_d = 258$ vertices) (**b**) low resolution case ($N_d = 92$ vertices)

Table 3 Benefits of perceptual haptic filters

	High direction res.		Low direction res.	
	CR	NMSE (dB)	CR	NMSE (dB)
Single frame case				
k=0.01	0.0669	−36.6375	0.0623	−33.3948
k=0.1	0.0260	−34.6063	0.0241	−31.3030
k=0.25	0.0138	−32.9836	0.0129	−29.7378
Multi frame case (m=2)				
k=0.01	0.0071	−33.8212	0.0063	−31.8455
k=0.1	0.0042	−32.7541	0.0039	−30.6527
k=0.25	0.0027	−31.8516	0.0024	−29.5160
Multi frame case (m=5)				
k=0.01	0.0047	−34.5251	0.0042	−31.9799
k=0.1	0.0020	−33.1267	0.0018	−30.2778
k=0.25	$9.5580e-04$	−31.7388	$8.8290e-04$	−28.4904

To evaluate the device filter and quantization effects of the compression approaches we executed the approach described above using the single frame depth image of Fig. 7. In Fig. 9a and b, we present the obtained NMSE assuming two different resolutions for the quantization of the directions of the generated force vectors. In each case we have considered different values for the haptic device spatial resolution (e.g., $n_w = 5, 6, 7$) and different amplitude quantization resolution ($n_q = 4, 6, 8$) for the generated force vectors determining the corresponding CR values presented in Fig. 9a and b respectively. By inspecting the derived results it is clearly shown that the achieved CR is lower than 5 % while the reconstruction quality is considerably high (e.g., the NMSE is less than −30 dB).

Moreover, in order to evaluate the effects of a perceptual Haptic filter in both CR and NMSE we provide Table 3 with results considering different values for the dead band parameter k for both the single and multi frame case. It is obvious that both the device and the perceptual filter introduced, reduce significantly the CR without affecting noticeably the reconstruction quality. In the multi-frame case the results are even more impressive since we exploit also the redundancy between consecutive frames. The interframe redundancy computes the difference between consecutive vectors corresponding to the 3D spaces, and only this difference is further processed as in the single frame case.

6 Discussion and conclusions

The proposed framework aims to propose a different way on how haptic rendering could be potentially dealt with, targeting at haptic media that can be processed, edited, modelled, indexed as a stand alone entity and not only as a result of the interaction between a user and the media space. It is noteworthy that the work towards this objective has revealed some very interesting limitations of the proposed scheme that turn into challenging areas for future work.

6.1 Dimensionality

The dimensionality of haptic perception is in the best case much higher with respect to visual perception as very well described in [6] and in the worst case infinite. Therefore, the transition of interaction-based rendering to open-loop stand-alone streaming haptic media, may introduce a significant increment in dimensionality resulting also in vast amount of haptic information to be transmitted and rendered if not dealt with explicitly. However, current haptic devices have limited degrees of freedom compared to the human perceptual capabilities. This unfortunate fact turns into an advantage as far as compression is concerned, since it significantly limits the amount of information that needs to be rendered as described in Section 4 (Type 1 vs Type 2 device). Moreover, typical solutions of the computer graphics research community could be employed, like space partitioning and culling. Another significant barrier is the fact that the probability mass function of the haptic medium, theoretically depends on the interaction. But we are trying to pull interaction out of the "equation"! Finally, to the author's view and as a major future research direction, a fundamental theoretical treatment of dimensionality reduction is necessary so as to result in a complete and tractable information theoretic framework of haptic media.

6.2 "Multi-view" haptic rendering

In Section 4 the effect of haptic filters over two haptic devices, namely Type 1 and Type 2 has been discussed. A direct equivalence of the Type 2 scheme in vision is single view rendering, which means that the scene is rendered without taking into account the viewers positions, which is the case in today's broadcast television. Similarly, a Type 2 haptic device renders the same haptic feedback no matter the position/orientation of the user. On contrary, a Type 1 haptic device, enables navigation of the user in the scene, since it encodes all the information of the force field. The equivalent model to vision, would be to transmit all images with respect to all potential positions/orientations of the user and render only the appropriate one. It is evident that this scheme encodes redundant information.

The computer vision society has partially tackled this problem by defining multi-view sequences and then synthetically and on-demand interpolating all intermediate views, based on specific assumptions. A similar formulation for haptic rendering is essential in order to minimize the dimensionality of the haptic media. In particular, this very challenging problem can be defined through the following question. *What is the minimum amount of points in space needed and what is their location, so that the full force field can be reconstructed with minimum error?* By answering the above question, under specific assumptions, e.g. on spatial frequency, estimation error, etc., a large step will have been taken towards the direction of compressed haptic media that can be navigated.

6.3 Closed-loop interaction and dynamic geometry

Another interesting aspect that has not been dealt with so far, is the possibility to alter, e.g. deform, the haptic media through interaction. I.e. the user not only navigates the haptic medium, but also modifies it, thus leading to highly realistic interactions, similarly to the traditional closed loop haptic rendering and interaction paradigm.

By definition, the proposed scheme is not able to cope with two-way interactions, since this would imply that we are closing the loop between the transmitter and the receiver. However, it has not necessarily to be the case. For example, free-form deformation of a haptic medium could be potentially performed, without closing the loop with the transmitter, by transmitting some control points together with the haptic medium and using a protocol on the how the medium is transformed based on the position of the control points. Moreover, assuming that a scene is described and rendered following the MPEG-4 paradigm, each object includes such metadata so that it knows "how to deform or transform itself". This way each object is assigned a haptic signal that transforms together with the underlying object. These approaches are however yet to be investigated in detail.

6.4 Vibrotactile and force feedback

The presented framework has dealt with the representation, rendering, coding of force vectors. Vibrotactile feedback, is however a very important and popular type of feedback that can be easily rendered using low-cost actuators. One of the major differences of an approach including vibrotactile feedback, is that, apart from the force magnitude, it has to deal also with the frequency of the vibration, since it is a very important perceptual cue. So an instance of a vibrotactile signal can hardly be defined for a specific time instance, but rather for a time window. An approach trying to merge force and vibrotactile feedback could potentially process the vibrotactile feedback in the spectral domain and then modulate the force feedback appropriately in the time domain. This concept has yet to be investigated as well.

References

1. Borst C (2005) Predictive coding for efficient host-device communication in a pneumatic force-feedback display. In: Proceedings first joint eurohaptics conference symposium. Haptic interfaces for virtual environ. Teleoperator system, pisa, Italy pp 596–599
2. Cha J, Ho YS, Kim Y, Ryu J (2009) A framework for haptic broadcasting. IEEE Multimedia 16(3):16–27
3. El-Saddik A, Orozco M, Asfaw Y, Shirmohammadi S, Adler A (2007) A novel biometric system for identification and verification of haptic users. IEEE Trans Instrum Meas 56(3):895–906. doi:10.1109/TIM.2006.887174
4. Guruswamy V, Lang J, Lee WS (2011) Iir filter models of haptic vibration textures. IEEE Trans Instrum Meas 60(1):93–103. doi:10.1109/TIM.2010.2065751
5. Hamam A, Saddik A (2013) Toward a mathematical model for quality of experience evaluation of haptic applications. IEEE Trans Instrum Meas PP(99):1–1. doi:10.1109/TIM.2013.2272859
6. Hayward V (2011) Is there a plenhaptic function. Phil Trans R Soc B 366:3115–3122
7. Hinterseer P, Hirche S, Chaudhuri S, Steinbach E, Buss M (2008) Perception-based data reduction and transmission of haptic data in telepresence and teleaction systems. IEEE Trans Signal Process 56(2):588–597
8. Hirche S, Bauer A, Buss M (2005) Transparency of haptic telepresence systems with constant time delay. In: Proceeding IEEE International Conference Control Appl., Toronto, Canada pp 328–333
9. Hossain S, Rahman A, El-Saddik A (2011) Measurements of multimodal approach to haptic interaction in second life interpersonal communication system. IEEE Trans Instrum Meas 60(11):3547–3558. doi:10.1109/TIM.2011.2161148

10. Jianxiong X, Owens A, Torralba A (2013) SUN3D: A Database of Big Spaces Reconstructed Using SfM and Object Labels. In: 2013 IEEE International Conference on Computer Vision (ICCV), pp 1625-1632
11. Kim M, Lee S, Choi S (2014) Saliency-driven real-time video-to-tactile translation. IEEE Transactions on Haptics 7(3):394–404
12. Kostopoulos K, Moustakas K, Tzovaras D, Nikolakis G, Thillou C, Gosselin B (2007) Haptic access to conventional 2d maps for the visually impaired. Springer International Journal on Multimodal User Interfaces 1(2):13–19
13. Kron A, Schmidt G, Petzold B, Zh MF, Hinterseer P, Steinbach E (2004) Disposal of explosive ordnances by use of a bimanual haptic telepresence system. In: Proceedings IEEE International Conference Robot. Autom., New Orleans, LA pp 1968–1973
14. Kuschel M, Cremer P, Buss M (2009) Passive haptic data-compression methods with perceptual coding for bilateral presence systems. IEEE Trans Syst Man Cybern Syst Hum 39(6):1142–1151
15. Laycock S, Day A (2007) A survey of haptic rendering techniques. Comput Graphics Forum 26(1):50–65
16. Lin M, Otaduy M (2008) Haptic rendering: foundations, algorithms and applications. p A.K.Peters publishing
17. Lloyd AS (2006) Least squares quantization in PCM. IEEE Trans Inf Theor 28(2):129–137. doi:10.1109/TIT.1982.1056489
18. Moustakas K (2013) Haptic media from an information theoretic perspective. In: IEEE International symposium on haptic, audiovisual environments and games, IEEE HAVE 2013, Istanbul
19. Moustakas K, Nikolakis G, Tzovaras D, Strintzis M (2005) Stereoscopic video generation based on efficient layered structure and motion estimation from a monoscopic image sequence. IEEE Trans Circuits Syst Video Technol 15(8):1065–1073
20. Moustakas K, Tzovaras D, Strintzis M (2007) Sq-map: Efficient layered collision detection and haptic rendering. IEEE Trans Vis Comput Graph 13(1):80–93
21. Moustakas K, Nikolakis G, Kostopoulos K, Tzovaras D, Strintzis M (2007) Haptic rendering of visual data for the visually impaired. IEEE Multimedia 14(1):62–72
22. Nikolakis G, Koutsonanos D, Daras P, Moustakas K, Tzovaras D, Strintzis M (2006) Haptic interaction in medical virtual environments. In: Wiley Encyclopedia of Biomedical Engineering, 6- Volume Set, Metin Akay (Editor) ISBN : 0-471-24967-X
23. Ortega A, Liu Y (2002) Lossy compression of haptic data. In: Touch in virtual environments: Haptics and the design of interactive systems. Englewood cliffs, NJ: Prentice-hall, ch. 6, pp 119–136
24. Ou E, Basdogan C (2002) Network considerations for a dynamic shared haptic environment. In: Proceedings Nat . Conf. Undergrad. Res Whitewater, WI
25. Robles-De-La-Torre G (2006) The importance of the sense of touch in virtual and real environments. IEEE Multimedia 13(3):24–30
26. Saddik E (2007) The potential of haptics technologies. IEEE Instrum Meas Mag 10(1):10–17
27. Sakr N, Georganas ND, Zhao J (2011) Human perception-based data reduction for haptic communication in six-dof telepresence systems. IEEE Trans Instrum Meas 60(11):3534–3546. doi:10.1109/TIM.2011.2161144
28. Shen X, Zhou J, El-Saddik A, Georganas ND (2004) Architecture and evaluation of tele-haptic environments. In: Eighth IEEE international symposium on distributed simulation and real-time applications, DS-RT, pp 53-60
29. Sakr N, Zhou J, Georganas ND, Zhao J (2009) Prediction-based haptic data reduction and transmission in telementoring systems. IEEE Trans Instrum Meas 58(5):1727–1736. doi:10.1109/TIM.2008.2009146
30. Srinivasan M, Basdogan C (1997) Haptics in virtual environments: taxonomy, research status and challenges. Comput Graph:393–404
31. Ternes D, MacLean K (2008) Designing large sets of haptic icons with rhythm. Springer LNCS Haptics: Perception Devices and Scenarios 5024:199–208
32. Vu MH, Na UJ (2011) A new 6-dof haptic device for teleoperation of 6-dof serial robots. IEEE Trans Instrum Meas 60(11):3510–3523. doi:10.1109/TIM.2011.2164285
33. Wang H, Liu X (2011) Haptic interaction for mobile assistive robots. IEEE Trans Instrum Meas 60(11):3501–3509. doi:10.1109/TIM.2011.2161141

Konstantinos Moustakas received the Diploma degree and the PhD in electrical and computer engineering from the Aristotle University of Thessaloniki, Greece, in 2003 and 2007 respectively. He has been a visiting lecturer in the same department during 2008. During 2003-2007 he served as a research associate in the Information Processing Laboratory of the Aristotle University of Thessaloniki and the Informatics and Telematics Institute, Centre for Research and Technology Hellas, where he has also served as a post-doctoral research fellow in the period 2007-2011. Since January 2012 he is an Assistant Professor in the Electrical and Computer Engineering Department of the University of Patras and Head of the Visualization and Virtual Reality Group. His main research interests include virtual augmented and mixed reality, haptics, virtual human modeling, information visualization, physics-based simulations, computational geometry, 3D content-based search, computer vision, and stereoscopic image processing. During the latest years, he has been the (co)author of more than 110 papers in refereed journals, edited books, and international conferences. He serves as a regular reviewer for several technical journals and has participated to more than 10 research and development projects funded by the EC and the Greek Secretariat of Research and Technology. He is a member of the IEEE, the IEEE Computer Society and Eurographics.

Aris S. Lalos received his Diploma degree, his M.A.Sc. degree and his Ph.D. from the Computer Engineering and Informatics Department (CEID), School of Engineering (SE), University of Patras (UoP), Rio Patras, Greece in 2003, 2005 and 2010, respectively. He has been a research fellow at Signal Processing and Communications Laboratory, CEID, SE, UoP, Rio-Patras, Greece from 2005 to 2010 and in Signal Theory and Communications (TSC) Department of the Technical University of Catalonia (UPC) from Oct. 2012-Dec. 2014. In period 2011-2012 was a telecommunication research engineer at Analogies S.A, an early stage start up. He is currently a postdoctoral researcher in the Visualization and Virtual Reality Group. His general research interest include, digital communications, adaptive filtering algorithms, wireless body area networks and biomedical Signal Processing. He is an author of more than 40 research papers in international journals, conferences and edited books. He has participated in several European projects related to the ICT and eHealth domain (e.g., COOPCOM, ALPHA, WSN4QoL, KinOptim, MyAirCoach etc.) and he acts as a regular reviewer for several technical journals. Aris Lalos received the best demo award in IEEE CAMAD 2014, while, in January 2015, he was nominated as Exemplary Reviewer for the IEEE Communications Letters.