Virtual skin: co-creating 3D materials with synesthetic artificial intelligence

Predrag K. Nikolić[1], Melanie McClintock[1], Giacomo Bertin[2], Jyotsna Parthasarathy[1]

[1]College For Creative Studies, Detroit, United States
pnikolic@collegeforcreativestudies.edu
mmcclintock@collegeforcreativestudies.edu
jparthasarathy@collegeforcreativestudies.edu
[2]Università degli Studi di Padova, Padua, Italy
giacomo.bertin@studenti.unipd.it

Abstract
Humans explore the world around us using our able senses. Numerous studies have suggested that the greater the number of sensory modalities stimulated at any one time, the richer our experiences will be, in that the experiences may have a higher level of impact and be more memorable across a variety of our senses. (e.g. Bahrick & Lickliter, 2000; Spence,2002; Stein & Meredith, 1993). This is becoming increasingly important and relevant as we enter the AI-generated virtual worlds where our visual perception should have a certain resemblance to our analog world experiences. If human perception is a cross-experiential phenomenon, how could we enrich artificial intelligence with multisensory creative abilities to support the same processes and outcomes for material interaction?

The ability to perceive correspondences across sensory modalities resembles the phenomenon of synaesthesia (Cytowic, 1989). Synesthesia is the combination of several senses simultaneously. It is an unusual condition which gives rise to a merging of the senses. For example, smells may trigger the experience of shapes, or letters may give rise to the perceptual experience of colour or may cause the experience of any combination of tastes, smells, shapes or sensations (Simner, 2012). These misunderstandings may be why the phenomenon was considered unreliable for over a century after its discovery by Francis Galton (1883; Sachs 1812). In the last decade, we can trace an increased interest in cross-sensory experiences signalling synaesthesia as its correlated phenomena (Dixon, Smilek, Cudahy, & Merikle, 2001; Mattingley, Rich, & Bradshaw, 2005; Ramachandran & Hubbard, 2001a, 2003). While for most, mental imagery may be most vivid in the case of visual images (Kosslyn, 1994), it is of keen interest to note that rich mental images can be generated via the sense of smell, touch, taste, and sound (e.g. Klatzky, Lederman and Matula, 1991; Reisberg, 1992; Stevenson and Case, 2005).

Through the intersection of two disciplines, User Experience Design (UXD) and Color and Materials Design (CMD), we aim to explore the possibilities of employing an intelligent agent to generate 3D materials based on merged multisensory inputs rather than relying on existing algorithms, visual styles, and previously constructed patterns. The exploration will be in the form of an experiment and workshop conducted with a trans-disciplinary group of graduate students from the User Experience Design and Colour and Material Design disciplines. We will use customised text to engage a 3D scene AI generator trained with available sensory data sets enriched with multi-perceptive input from the workshop participants. Our key novelty is the use of artificial intelligence trained with multi-sensory material properties to generate AI-generated 3D materials for virtual and potentially hybrid products and environments.

Author keywords
AI creativity, 3D virtual materials, Human-AI interaction, Synesthesia, Multi-sensory User Experience Design, Transdisciplinary Co-creation

Introduction
In this paper, we aim to investigate opportunities for involving artificial intelligence in the design process of creating 3D virtual worlds. Our focus will be on the AI generation of 3D digital materials and their potential to arouse sufficient sensory and emotional stimuli. Several studies have suggested that the greater the number of sensory modalities stimulated at any one time, the richer our experiences will be, in that the experiences may have a higher level of impact and be more memorable across a variety of our senses. (e.g. Nikolic & Russo, 2019; Bahrick & Lickliter, 2000; Spence,2002; Stein & Meredith, 1993). Accordingly, In the AI-generated virtual worlds, our sensory and cognitive expectations lead us to search for resemblance from our analogue world experiences. Hence, if human perception is a cross-experiential phenomenon, how could we enrich artificial intelligence with multisensory creative abilities to support the same processes and outcomes for material interaction?

Through the intersection of two disciplines, User Experience Design (UXD) and Color and Materials Design (CMD), we aim to explore the possibilities of employing an intelligent agent to generate 3D virtual materials based on merged multisensory inputs rather than relying on existing algorithms, visual styles, and previously constructed patterns.

Our approach is hybrid and includes several methods applied in different research stages. In the first phase, we conducted a single case experiment, which we will describe in detail and present the findings. This led us to the second phase, where we prototyped and experimented with the AI-driven text-to-image platform for generating and rendering virtual materials on 3D objects. The whole process and the framework used are described in this paper's Synesthetic Artificial Intelligence Generation chapter. Lastly, we plan to conduct contextual interviews in a workshop with a mixed group of students selected from the User Experience Design (UXD) and Color and Materials Design (CMD) majors.

Our key innovation is using artificial intelligence trained with multi-sensory material properties to generate AI-generated 3D virtual materials for meta – and potentially hybrid – products and environments.

Background

In our conceptual approach, synesthesia plays an important role in designing multi-sensory meta-perception beyond the physical world. Synesthesia is a psychological phenomenon related to the human brain's capability to combine several senses simultaneously. In particular, stimuli on one particular physical sensory can produce additional sensory experiences for which sensory inputs do not exist (van Leeuwen, Singer & Nikolić, 2015). It is an unusual condition that gives rise to a merging of the senses. For example, smells may trigger the experience of shapes, or letters may give rise to the perceptual experience of colour or may cause the experience of any combination of tastes, smells, shapes, or sensations (Simner, 2021). One of the most reported types of synesthesia is audio-visual (AVS)(Afra, Anderson, Funke et al, 2012), where sound stimuli can evoke the visual experience and vice versa. In particular, sounds such as the middle note 'C' can induce a red colour experience, but the same note three octaves higher can become green (Ginsberg, 1923). Likewise, we can have different varieties of sense responses to sensory stimulation, such as; the taste of a particular food (gustatory) can associate us with the visual appearance of the food (Cytowic, 1989); when we hear a particular sound (auditory) can induce smell of a specific food (olfactory) (Beeli, Esslen & Jäncke, 2006). Therefore, despite early scepticism, contemporary research in the field proves that synesthesia exists.

Furthermore, the interaction with materials in regard to a product, space, or environment is essential for a completely holistic multi-sensory user experience. Hence, from the perspective of colour and material design, "texture is an essential component of materiality that evokes the sense of tactility," often translated through a flat, 2D image that is applied onto the surface of a 3D object to define its colour, gloss, transparency, roughness, and other key physical attributes (Lefteri, 2014). A texture in 3D rendering engines like Photoshop, Substance Painter, Substance Designer, Blender, and Keyshot require inputs whose details must be specified unambiguously. These tools make use of generating textures in a *procedural material workflow*. (Jantunen, 2017). Procedural materials are created using mathematical models instead of pre-stored data, reducing the storage needed for these materials and thus creating texture maps to be applied to models in an actual runtime. Dong et al., 2019).

As an example case study, *Procedural Material Generation* (Jantunen, 2017) delves into the steps to create a procedural material in five steps with a focus on different material properties such as texture, patterns, micro-surfaces, colour, and lastly, defines the *procedural aspect* of the material to identify the properties which can be changed while still retaining the essence of the material. Our *workflow* (Figure 1) combines the aspects of synesthetic modalities with procedural material workflows to generate multi-sensorial materials, unlike a 3D rendering engine, whose inputs must be specified unambiguously and in complete detail, without further sensory extension when it comes to no explicitly stated properties (Ramesh, et al., 2021). The design of products and experiences are typically the result of many fine-meshed and complex processes (Jongerius, 2010); and our proposed workflow uses the sensory nature of synesthetic procedurality when designing multi-sensory meta-perceptions.

This multi-sensory, layered approach is based on an archival collection of information sourced through 1) scent, 2) sound, 3) emotion, 4) taste, and 5) movement. The sensory experiences are categorised as layers that, when combined together, compose a unique material outcome.

ARCHIVE OF INFORMATION

LAYER 1	LAYER 2	LAYER 3	LAYER 4	LAYER 5
SCENT	SOUND	EMOTION	TASTE	MOVEMENT

Figure 1. Multi-sensory layered approach

During the first phase, digital prototypes of 3D virtual materials were created using Open AI's text-image generation AI system DALL.E. Over 150 unique samples were created through prompts that described various material attributes regarding colour, texture, finish, and form. For the experiment, fifteen samples of visual distinction were short-listed (Figure 2).

Figure 2. Virtual material samples created on DALL.E with prompts (left to right) beige green exploding fuzzy fluidic sphere, black white eternal colliding amorphous sphere, and high gloss stretchy blob slime beige sphere

Experiment

A single-case experiment research was employed to initiate the first phase of the 3D virtual materials generation. We invited six graduate students to participate in the experiment from two graduate design programs, User Experience Design (three participants) and Color and Materials Design (three participants). The participants included three females, two males, and one non-binary, of American, Indian and Mexican origins, and in an age range of 23-27 years old.

The experiment was conducted in a controlled lab environment with a defined layout, ambient temperature, sound, light, and projected image control (Figure 3). The AI-generated digital material samples were displayed on the screen for a period of three minutes, during which the participants observed the sample and noted their sensorial responses in a provided worksheet. (Table 1). The experiment lasted 45 minutes, followed by 15 minutes of guided discussion.

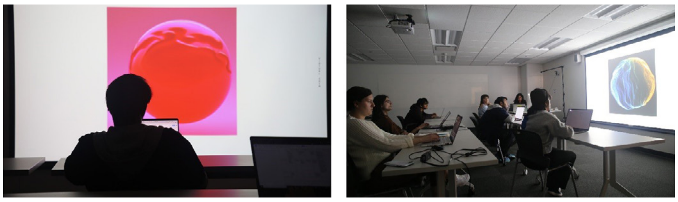
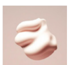


Figure 3. Single-case experiment set-up

Findings

Upon the completion of the first phase, participants revealed directives toward sensory modalities in regard to a series of AI-generated materials. The results from the experiment revealed the ability of the participants to look beyond the aesthetic attributes of the virtual materials. With this in consideration, it should be noted that half of the participants stated had it not been for the structure of the worksheet, they would not have related the visual with all of the senses. Four out of six participants stated that they found associations to be made quicker with certain senses than others. The sensorial responses experienced by the participants were then used as inputs to generate the synesthetic AI-generated virtual skin.

Table 1. Example of worksheet and responses from the experiment

	Touch	Smell	Taste	Sound	Visual	Emotions
Sample 1 	Tender Soft	Soft scent of a mother	Delicate sweetness	Lullaby	Baby's skin	When pressed will release the flow of milk
Sample 2 	Fuzzy Soft	Wet grass	Fresh taste of blooming spring time	Soft breeze	A planet with unknown biome	A walk through paddy fields in the morning

Synesthetic Artificial Intelligence Generation

Starting from text descriptions, we used DALL-E to generate the image of a sphere texturized according to the initial prompt – as demonstrated in the single-case experiment – and then we extracted the corresponding texture. The procedure to fit the bump map and the texture was then refined using the Stable Diffusion model to progressively refine the starting template mesh, which in our case, is a sphere.

A neural network generates an RGB texture and predicts the magnitude of the movement along the normal direction for each vertex. To compute the displacement, we encode the position of the vertex $x \in R^{Nx3}$ using Multiresolution Hash Encoding (MHE). We parsed the features obtained $h \in R^{NxF}$ to a SirenNet, an MLP (MultiLayer Perceptron) with sinusoidal activation function:

$$\bar{x} = \bar{x} + \bar{n} \cdot d \cdot SirenNet(MHE(\bar{x}))$$

where $\bar{n} \in R^{Nx3}$ are the vertices normal and d the maximum displacement. The texture is computed in a similar way by mapping the positions of the pixels to the range $[-1,+1]$,

$$RGB_{i,j} = SirenNet\,(\sin \sin ((2i/(H-1)-1)\,\pi), \sin \sin ((2j/(W-1)-1)\pi))$$

where H, W are the texture dimensions. This initial transformation ensures periodical textures. Using a sinusoidal activation function makes the training shorter and more stable for this task. During the generation, we optimise the material gloss level to reduce artefacts, like light reflection included in the texture.

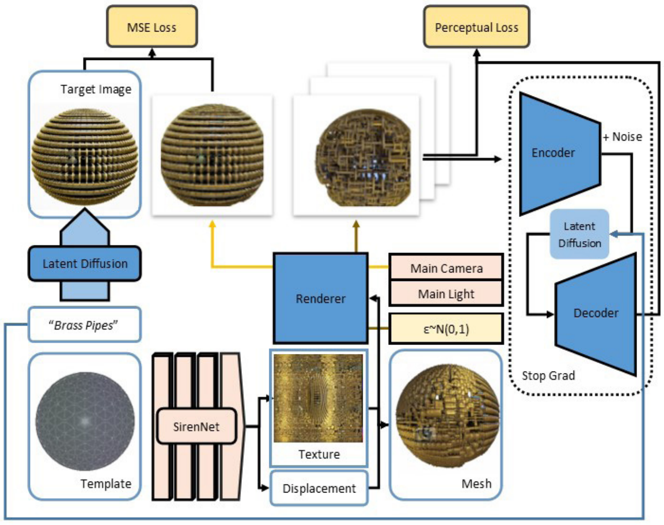


Figure 4. AI Synesthetic Virtual Materials Generation Procedure

For each training step, we start with generating the texture and the displacement along the normal for each vertex, then we produce a refined version of each render using the Stable Diffusion Img2Img pipeline conditioned on the text prompt. Finally, we compute the MSE (Mean Square Error) between the main render and the target image, then compare the renders from the cameras sampled uniformly around the sphere and their refined version using a perceptual loss based on a pre-trained VGG19 net.

This solution led to faster convergence and better quality than using a loss based on CLIP embeddings, moreover the refining procedure must be executed without the gradient computation, allowing to reduce the memory occupancy and increasing the training speed if compared to the SDS (Score Distillation Sampling) used in Magic3D, because it does not require backpropagation through the encoder network at the cost of having to decode the denoised latent vector. In this way, we produce textures with 712 pixels of resolution. To enforce a more coherent bump map, we follow the same procedure with a uniform texture. To guarantee stability during the training, we must ensure semantic consistency between the refined images and the corresponding renderer. For this reason, we set a low strength (around 20%) and a not too high guidance scale (we used 12.5).

Conclusion

We aim to explore the possibilities of using AI for 3D virtual material generation, with the potential to arouse sufficient sensory and emotional stimuli. Hence, we have collected and designed multi-sensory material properties to generate AI-generated 3D materials for virtual – and potentially hybrid – application to products and environments through User Experience Design and Color and Material Design (CMF) workflows.

In the first of several on-going experiments, we have investigated procedural approaches to collect sensory data to train our neural network model toward enabling its synesthetic reasoning and virtual material rendering based on multi-sensory data inputs. By increasing the number of sensory modalities stimulated in the multi-layered, procedural material approach, the sensorial experiences related by the observer have an increased impact in regard to their virtual material experience.

While we seek to find a balance between our real and virtual worlds, in turn, the expectations of our senses have linked our findings to human perception within the collaborative environment where User Experience Design and Color and Material Design enrich the ability to support synesthetic material interaction.

References

Afra P, Anderson J, Funke M et al. (2012) Neurophysiological investigation of idiopathic acquired auditory–visual synesthesia. Neurocase 18(4), pp. 323–329.

Bahrick LE, Lickliter R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. Developmental psychology 36(2), pp. 190.

Beeli, G., M. Esslen, and L. Jäncke. (2005). Synaesthesia: when coloured sounds taste sweet, Nature (434:7029), pp. 38-38.

Cytowic, R. E. (1993). The man who tasted shapes: A bizarre medical mystery offers revolutionary insights into reasoning, emotion, and consciousness, New York: Putman.

Dinh, H. Q., Walker, N., Song, C., Kobayashi, A. and Hodges, L. F. (1999, 13–17 March). Evaluating theimportance of multi-sensory input on memory and the sense of presence in virtual environments. Paper presented at the IEEE Virtual Reality Conference, Houston, TX.

Dong, J., Wang, L., Liu, J., Gao, Y., Qi, L., & Sun, X. (2019). A procedural texture generation framework based on semantic descriptions. *Knowledge-Based Systems, 163*, 898–906. https://doi.org/10.1016/j.knosys.2018.10.014

Ginsberg, L. (1923) A case of synaesthesia. The American Journal of Psychology, pp. 582-589.

Hoffman, H. G., Hollander, A., Schroder, K., Rousseau, S. and Furness, T. I. (1998). Physically touching and tasting virtual objects enhances the realism of virtual experiences. Journal of Virtual Reality, 3, 226–234.

Jantunen, J. (2017). "Creating Procedural Textures for Games with Substance Designer"

Klatzky, R. L., Lederman, S. J. and Matula, D. E. (1993). Haptic exploration in the presence of vision. Journal of Experimental Psychology: Human Perception and Performance, 19, 726–743.

Kosslyn, S. M. (1994). Image and brain: the resolution of the imagery debate. Cambridge, MA: MIT Press.

Lefteri, C. (2014). *Materials for Design*, London: Laurence King Publishing.

Lindstrom, M. (2005). Brand sense: build powerful brands through touch, taste, smell, sight, and sound. New York: Free Press.

Ramesh, A., Pavlov, M., Goh, G., Gray, S. (2021, January 5). DALL.E: Creating Images from Text. *OpenAI*. https://openai.com/blog/authors/mikhail/

Schouwenberg, L., Antonelli, P., & Rawsthorn, A. (2011). *Hella Jongerius: Misfit*. Phaidon Press.

Schifferstein, H. N. J, Spence, C. Multisensory Product Experience, Ed. Schifferstein, H. N. J. Ed. Hekkert, P. (2007), Product Experience, Elsevier Science.

Simner J. (2012) Defining synaesthesia. British journal of psychology 103(1), pp. 1–15.

Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2001). Synaesthetic photisms influence visual perception. Journal of Cognitive Neuroscience, 13, 930–936.

Spence, C. (2002). The ICI reports on the secret of the senses. London: The Communication Group.

Spence, C. (2007a). Audiovisual multisensory integration. Acoustical Science and Technology, 28, 61–70.

Stein, B. E. and Meredith, M. A. (1993). The merging of the senses. Cambridge MA: MIT press.

Ramachandran VS, Hubbard EM. (2003). The phenomenology of synaesthesia. Journal of consciousness studies 10(8), pp. 49–57.

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: Implications for the role of early experience in lexical–colour associations. Cognition, 98, 53–84.

van Leeuwen, T.M., Singer, W. and Nikolić, D. (2015). The merit of synesthesia for consciousness research. Frontiers in psychology, 6, pp.1850.

Washburn, D. A., Jones, L. M., Vijaya Satya, R., Bowers, C. A. and Cortes, A. (2003). Olfactory use in virtual environment training. Modelling and Simulation, 2(3), 19–25.

Whipple G.M. (1900). Two cases of synaesthesia. American Journal of Psychology, 11: 377-404,