ACM/EG Expressive Symposium
Eurographics Workshop on Intelligent Cinematography and Editing (2025), pp. 1–4
D. Berio and A. Bruckert (Editors)

Evaluating Temporal Coherence using a Watercolor Renderer

I.E.C. Morgan[1] and M. Billeter[1] and K. Anjyo[2] and RK. dos Anjos[1]

[1]University of Leeds, School of Computer Science, United Kingdom
[2]OLM Digital, Japan

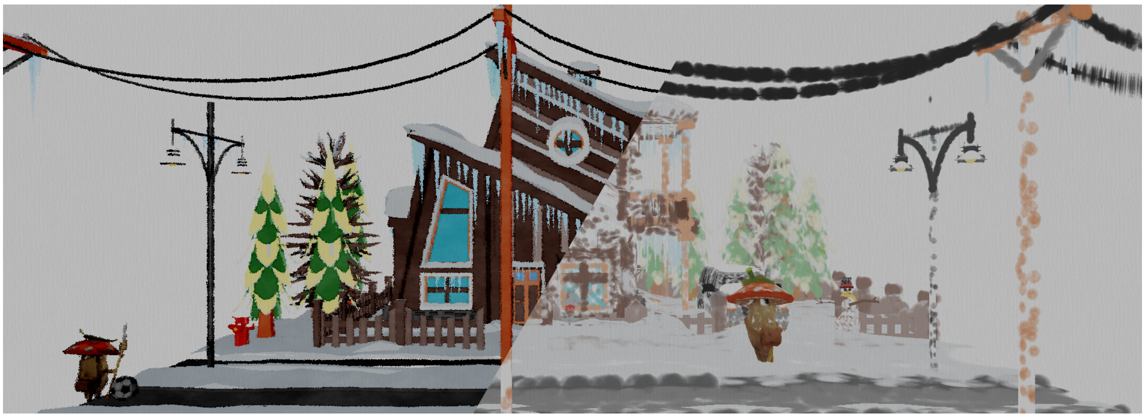


Figure 1: *Two different styles visible in our watercolour renderer: a screen-space approach [LMHB00], and an object space approach [Mei96].*

Abstract

Temporal coherence is a long-standing issue within Non-Photorealistic Rendering (NPR). The problem has been defined as a trade-off between the three main factors: flatness, temporal continuity and motion coherence. Approaches that improve temporal coherence are applied across different styles within diverse animation contexts. We have implemented a watercolour renderer that supports multiple temporal coherence approaches in a unified system to investigate this trade-off. The approaches are then evaluated against existing work, with consideration of how external factors including animations and textures may influence perceived incoherence.

CCS Concepts

• ***Computing methodologies*** → ***Non-photorealistic rendering;*** *Animation; Perception;*

1. Introduction

Non-Photorealistic Rendering (NPR) refers to rendering techniques that target visual styles beyond physical realism. It has proved itself a valuable tool in the film and video game industries to help tell stories and distinguish new work. To replicate characteristics of traditional art, random generation is used to procedurally create features like brush strokes and pigment effects but can cause distracting temporal artifacts when animated. This problem is related to *temporal coherence*, specifically to the consistency of stylization features and their movement between frames.

Bénard et al. [BBT11] defines temporal coherence as a problem of balancing three factors: motion coherence, temporal continuity and flatness. In this poster we refer to this as the *temporal coherence trade-off*. Notably, they illustrated how these three components are contradictory, and how different combinations of each induce specific artifacts (Figure 2).

While there have been several works trying to address this issue, they usually target different artistic styles where different aspects of the trade-off have higher relevance. As a result, individual systems are not conducive to comparative studies of factors or unified explorations of the trade-off.

To establish a unified scope of the temporal coherence trade-off and its comparative study, we have implemented eight different approaches based on six published methods [BKTS06, BBT09, KP11,

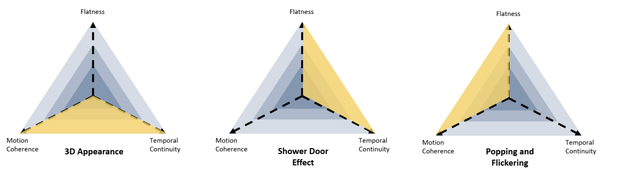


Figure 2: *The three different extremes of temporal incoherence presented by Bénard et al [BBT11] alongside the visual artifacts they produce, referred to as the 'temporal coherence trade-off'*

BSM*07, LMHB00, Mei96] that address aspects of temporal coherence trade-off. Our implementation uses a simple watercolour renderer to verify the impact of this trade-off in a constrained scenario. Using an interactive demo, we discuss the extent of its impact to the perceived quality of the result.

2. The Temporal Coherence Trade-Off

Popping and flickering refers to how primitives and artifacts used in stylization appear and disappear discontinuously between frames. Initial work has considered this undesirable due to how the human visual system (HVS) is highly sensitive to flicker, particularly in the near periphery [SHB*17], hence these artifacts can unduly attract attention to an unintended region. High amounts of flickering can typically be related to motion coherence, as primitives attach themselves to the underlying motion of a surface, which may move so quickly that strokes 'scintillate'. Alternatively, popping and flickering can be caused by viewpoint changes, especially in 3D where theses changes also affect silhouettes. Errors in primitive placements at silhouettes can lead to strokes inconsistently appearing and disappearing resulting in popping and flickering.

To avoid popping and flickering many techniques will make marks or textures static, ensuring they never change between frames and in turn increasing temporal continuity. This induces the shower door effect, where the stylization has no correlation to scene motion. In highly animated scenes, incorrect motion perception may have a more adverse effect than flickering.

Flatness is the perception of a lack of three-dimensionality in the rendered objects. It relates to the other two aspects in the fact that in the absence of binocular disparity motion becomes an important component of our depth perception. In NPR animations, flatness is targeted due to its appearance in traditional hand-drawn animations where flatness typically arises from the drawing process itself. To avoid a 3D appearance, previous work has focused on ensuring stylization elements are disconnected from the underlying geometry (e.g. maintain size with depth) or by ensuring marks are placed screen space.

3. Prototype

We adapt the presented techniques to fit a naive watercolour rendering implementation in a real time renderer, allowing control of different parameters from each method. We chose watercolour rendering as it has a wide range of acceptable representations contemporary media, and is simple to implement. This means we are able

Figure 3: *Apple [Tah20] and Mushroom Wonderland scenes [MV20] [Edw19] [BSl17].*

to meaningfully adapt work that was done in other styles. The renderer was implemented using the OpenGL API with a simple UI for controlling parameters.

The implementation is based on the work by Bosseau et al. [BKTS06]. 3D objects are rendered using a cel-shader and the colour modification model introduced by Bosseau et al. [BKTS06] to simulate limited colour palettes. Edges are detected using a simple Sobel filter, and noise is generated [KP11], retrieved from 2D textures [BSM*07, CDLK20, LMHB00] or from 3D textures [BBT09] to replicate pigment density effects[†].

Edge wobbling is applied using substrate texture information [MCS19] and a global strength parameter. Both pigment effects and darkened edges are applied with the colour modification model before the substrate is applied to the entire result in screen space. The key difference in the implemented approaches is the source or manipulation of this noise, which will be discussed in the following section.

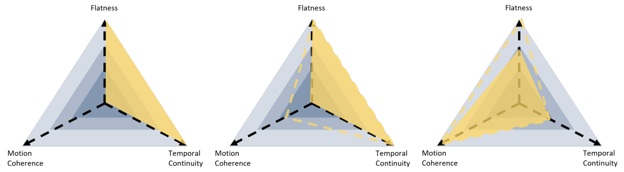
4. Evaluation

We evaluated all eight approaches, further detailed in Sections 4.1-4.3, using the two scenes depicted in Figure 3. The the *Apple* scene contains a single object with a realistic texture and all possible movements animated in a loop (image plane translation, depth translation, rotation in place). The *Mushroom Wonderland* scene depicts more complex motion (mushroom characters waddling) and parallax cues, but overall simpler objects with mostly single coloured textures. All of the following interactions can be verified through our demo, or supplemental video.

4.1. Flat and Temporally Continuous.

For the initial approach, we map the noise texture with screenspace coordinates following the work of Lake et al. [LMHB00]. This is the definition of a shower door effect (Figure 2 center), and it is easily verifiable that it preserves the feeling of flatness and temporal continuity (Figure 4a). The approach from Breslav et. al [BSM*07] has the goal of reducing the shower door effect by moving the texture in screenspace based on the movement object samples in 3D (a similar approach has also been applied to watercolour [BKTS06]). We find that it does indeed preserve the flatness, and increases the

† 2D textures are from `https://opengameart.org/content/700-noise-textures` and 3D textures are from `https://github.com/Calinou/free-blue-noise-textures/tree/master`

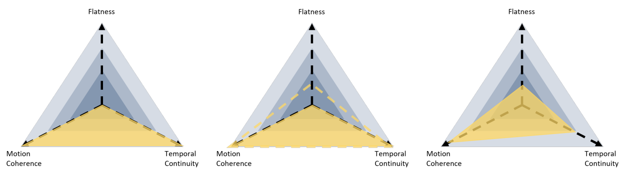


(a) *[LMHB00]* **(b)** *[BSM*07]* **(c)** *[KP11] with* $\alpha \approx 1$

Figure 4: *Trade offs for flat and temporally coherent techniques.*

motion coherence for translating motions. However, for rotations in the *Apple* scene we found that shower door was still visible. When combined with detailed textures, the combination of flat noise and perspective-distorted apple texture exaggerated the shower door effect. On more complex motions, we found that the watercolour pattern tends to still be incoherent with motion, which may have an effect that is visually worse than showerdoor, as it gives a different motion cue than the desired one. Thus we consider the increase in motion coherence to be motion dependent for this second approach (Figure 4b). The work from Kass and Pesare [KP11] also aims to remove showerdoor effect by tracking noise samples according to the scene's motion. We notice popping was still present in disocclusions, where new noise samples need to be introduced (Figure 4c). In this approach, it is simple to achieve total motion coherence by increasing the α parameter which increases the amount of frames before full regeneration of noise. However motion coherent noise samples gave the perception of a 3D appearance in the *Apple* scene, as their movement was easy to track.

4.2. Continuous and Coherent: 3D



(a) *Noise textured object space* **(b)** *[Mei96] in object space* **(c)** *[Mei96] in screen space*

Figure 5: *Trade-offs for continuous and coherent techniques.*

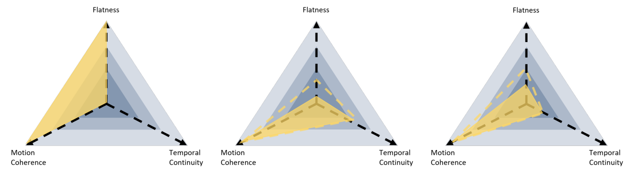
If flatness is not a goal, a watercolour renderer can be achieved by mapping our noise texture to the object, using its own texture coordinates (Figure 5a). While this may sound like a naive approach, it is in fact what happens in cases where artists manually paint textures on objects, with the distinction that our textures are considerably simpler.

In order to have more flexibility in our watercolour style representation and explore the tradeoffs existing in an object space approach, we adapted our renderer to perform splatting [Mei96], which is a common paradigm in painterly rendering. The cel-shading pass was replaced by generating a controllable number of splats per triangle using a geometry shader. Every triangle in the mesh will generate a random distribution of splats that are ensured to stay the same per frame, as splats are always generated based on the same seed values. For our brush textures we randomly select from three noise textures [Stu21] with circular faded edges using the same consistent per-triangle values. We compare two alternative splatting approaches; object oriented and screen space splats.

We find that for the object-space splats, we could still perceive a significant level of flatness in the *Apple* scene, which was contrary to what can be noticed in similar painterly work [dARLP18]. The perceived flatness here was considerably higher than the naive object texturing approach. We attribute this to the fact that the splatting approach removes high frequency patterns present in the realistic apple texture, which serve as a depth cue that can be tracked in rotating motion. Moreover, with soft edged splats used to replicate the watercolour style (differently from typical painterly approaches), individual splats are harder to track with object motion, contributing to the flat appearance (Figure 5b)

When splats are oriented towards the screen, we could notice a flatter appearance, as they are of equal size throughout the whole scene. This comes with the side-effect of popping artifacts being significantly more visible in rotating motions (Figure 5c).

4.3. Flat and Motion Coherent



(a) *[KP11] with* $\alpha = 0$ **(b)** *[BBT09]* **(c)** *[CDLK20]*

Figure 6: *Trade-offs for flat and motion coherent techniques.*

Finally, if we consider motion coherence to be highly desirable, the tradeoffs would be reducing the perceived flatness to increase temporal continuity, or the opposite. If we consider a naive approach of having a new noise texture at every frame, it would not be temporally consistent, but it could be considered motion coherent, as the noise representing watercolour splats does not serve as a cue for motion due to their randomness. In fact, this is what the approach from Kass et al. [KP11] without tracking samples would look like (Figure 6a). Given the considerations in Section 4.1, we find that controlling α is great example of a tradeoff between motion coherence and temporal continuity without influencing flatness.

In cases where objects are moving closer and away from the screen, fractalization approaches that blend different noise scales can reduce flickering, thus enforcing temporal continuity, and not sacrificing motion coherence. The work from Bénard on dynamic solid textures [BBT09] is one example. In our implementation we noticed that they achieve a flat look for objects in movements along the viewing direction, as the fractalization approach removes motion cues, however, they do not adress the 3D look from rotating motions. This, interestingly, is the opposite of what was noticed in the work from Breslav et al. [BSM*07] (discussed in Section 4.1). The work on meta textures [CDLK20] adresses this in a similar way, but by using a 2D texture instead of a 3D one, they achieve a flatter look in translating motions. However, there are increased popping and flickering artefacts depending on the underlying geometry.

5. Conclusion

From the evaluated techniques, we find that there is indeed evidence of a perceived trade-off between those three goals; flatness, temporal continuity and motion coherence. However, the effectiveness of different techniques is closely related to motion and texture content of the visualized objects. Particularly for rotations, Coherent noise [KP11] and Dynamic solid textures [BBT09] fail to keep flatness on rotating motions. Dynamic 2D patterns [BSM*07] is able to preserve it, but creating a shower door effect in the same scenario. Future work should look into combining these factors using higher level understanding of the rendered scenes.

There are, however, two other considerations to raise from our evaluation. Firstly, both Motion coherence and Temporal continuity can be used in NPR work as a stylistic choice and increase the perceived quality of the work (Section 4.3), as it may resemble hand drawn artifacts, while smooth and accurate motions can be easily identifiable as 3D-driven graphics. This is also verified by seeing numerous artistic projects that implement both popping and flickering, and the shower door effect as a stylistic choice. (e.g. Gankutsuou with screen-space textures, and Loving Vincent, a hand-painted animated movie).

Secondly we notice that style has an influence on perceived flatness with some styles like line drawing inherently appearing more flat. This has lead to improved results that can focus on controlling only motion coherence and temporal continuity [CS16].

Regarding flatness, we see from the splatting example [Mei96] that the perception of flatness is closely related to the sampling rate and texture content of the visualized object. Additionally, the realistic textured apple scene and Breslav et al.'s [BSM*07] dynamic 2D patterns generated an uncanny mix of 3D texture cues and flatness under rotation. These exemplify that although flatness interacts with the trade-off, it is mostly connected to the visual style and human perception of the rendered object. Further understanding on what operates as a depth cue for the HVS may allow decoupling of this factor from the other two.

In summary, we believe that these are factors that surely interact with each other given current approaches and implementations of NPR algorithms. However, our experiments suggest that considering combinations of human perception and controllable stylistic choices may lead to new potential avenues of research within temporal coherence in NPR.

References

[BBT09] BÉNARD P., BOUSSEAU A., THOLLOT J.: Dynamic solid textures for real-time coherent stylization. In *Proceedings of the 2009 symposium on Interactive 3D graphics and games* (Boston Massachusetts, Feb. 2009), ACM, pp. 121–127. URL: https://dl.acm.org/doi/10.1145/1507149.1507169, doi:10.1145/1507149.1507169. 1, 2, 3, 4

[BBT11] BÉNARD P., BOUSSEAU A., THOLLOT J.: State-of-the-art report on temporal coherence for stylized animations. In *Computer Graphics Forum* (2011), vol. 30, Wiley Online Library, pp. 2367–2386. 1, 2

[BKTS06] BOUSSEAU A., KAPLAN M., THOLLOT J., SILLION F. X.: Interactive watercolor rendering with temporal coherence and abstraction. In *Proceedings of the 4th international symposium on Non-photorealistic animation and rendering* (Annecy France, June 2006), ACM, pp. 141–149. URL: https://dl.acm.org/doi/10.1145/1124728.1124751, doi:10.1145/1124728.1124751. 1, 2

[BSl17] BSLEGT: Football - Download Free 3D model by BSlegt, Nov. 2017. URL: https://sketchfab.com/models/072a7e8d610f4143aa510256112d97cf/embed?autostart=1. 2

[BSM*07] BRESLAV S., SZERSZEN K., MARKOSIAN L., BARLA P., THOLLOT J.: Dynamic 2D Patterns for Shading 3D Scenes. 1, 2, 3, 4

[CDLK20] CURTIS C., DART K., LATZKO T., KAHRS J.: Real-time non-photorealistic animation for immersive storytelling in "Age of Sail". *Graphics and Visual Computing 3* (June 2020), 100012. URL: https://linkinghub.elsevier.com/retrieve/pii/S2590148619300123, doi:10.1016/j.cagx.2019.100012. 2, 3

[CS16] CARDONA L., SAITO S.: Temporally Coherent and Artistically Intended Stylization of Feature Lines Extracted from 3D Models. *Computer Graphics Forum 35*, 7 (2016), 137–146. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13011, doi:10.1111/cgf.13011. 4

[dARLP18] DOS ANJOS R. K., RIBEIRO C. S., LOPES D. S., PEREIRA J. M.: Stroke-based splatting: an efficient multi-resolution point cloud visualization technique. *The Visual Computer 34* (2018), 1383–1397. 3

[Edw19] EDWIIXGG: Low Poly Winter Scene - Download Free 3D model by EdwiixGG (@Edwin3D), Apr. 2019. URL: https://sketchfab.com/models/2ffe1ed449414cc184078dd1e2882284/embed?autostart=1. 2

[KP11] KASS M., PESARE D.: Coherent noise for non-photorealistic rendering. In *ACM SIGGRAPH 2011 papers* (Vancouver British Columbia Canada, July 2011), ACM, pp. 1–6. URL: https://dl.acm.org/doi/10.1145/1964921.1964925, doi:10.1145/1964921.1964925. 1, 2, 3, 4

[LMHB00] LAKE A., MARSHALL C., HARRIS M., BLACKSTEIN M.: Stylized rendering techniques for scalable real-time 3D animation. In *Proceedings of the 1st international symposium on Non-photorealistic animation and rendering* (Annecy France, June 2000), ACM, pp. 13–20. URL: https://dl.acm.org/doi/10.1145/340916.340918, doi:10.1145/340916.340918. 1, 2, 3

[MCS19] MONTESDEOCA S. E., CHIEW Y. X., SEMMO A.: Substrate profiles for MNPR, May 2019. URL: https://researchdata.ntu.edu.sg/dataset.xhtml?persistentId=doi:10.21979/N9/HI7GT7, doi:10.21979/N9/HI7GT7. 2

[Mei96] MEIER B. J.: Painterly rendering for animation. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (Aug. 1996), ACM, pp. 477–484. URL: https://dl.acm.org/doi/10.1145/237170.237288, doi:10.1145/237170.237288. 1, 3, 4

[MV20] MEOW-VORTEX: Mushroom Warrior - Download Free 3D model by Meow-Vortex, Apr. 2020. URL: https://sketchfab.com/models/f0ab1da8d7e74759b5bea8618c350914/embed?autostart=1. 2

[SHB*17] SINHA R., HOON M., BAUDIN J., OKAWA H., WONG R. O., RIEKE F.: Cellular and circuit mechanisms shaping the perceptual properties of the primate fovea. *Cell 168*, 3 (2017), 413–426. 2

[Stu21] STUDIOS S. B.: 700+ Noise Textures, July 2021. Publisher: OpenGameArt.org. URL: https://opengameart.org/content/700-noise-textures. 3

[Tah20] TAHIRNILIN: Apple - Download Free 3D model by TahirNilin, Jan. 2020. URL: https://sketchfab.com/models/643eb66651864bb78871e5c1066b4ef6/embed?autostart=1. 2