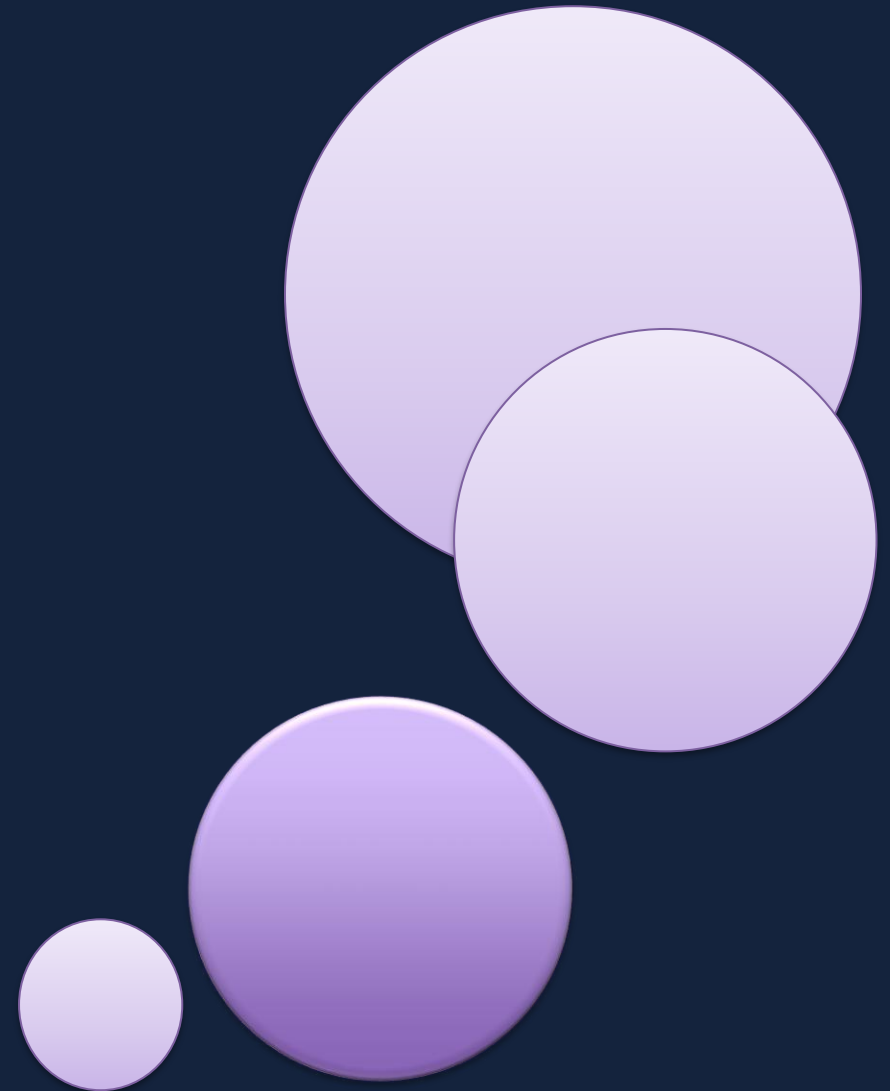


Cultural Representation in AI-Generated Visual Storytelling

Effects of Culturally Grounded Narrative Prompting

Aswathy Thoniyil Ravindran

Aswathy Thoniyil Ravindran,
PhD Student, Dongseo University, South Korea
Song Balgum, Associate Professor, Dongseo University, South Korea
ICCT 2026



Why cultural representation needs more than visual objects

A scene can contain culturally recognisable objects and still fail to represent the culture meaningfully.

- Text-to-video systems can produce realistic, coherent-looking sequences.
- Cultural details may be inaccurate, omitted, or disconnected from the scene.
- A narrative sequence raises an additional question: are cultural details integrated across events?

VISUAL PRESENCE

Are the expected objects, clothing, and practices visible?

CULTURAL CONTEXT

Do those elements fit the social and physical setting?

NARRATIVE INTEGRATION

Do cultural elements make sense within the story's progression?

Research gap and question

Prompt formulation

Prior work shows that prompt wording can influence cultural content in generated images.

Video evaluation

Benchmarks identify cultural failures in generated video, including fine-grained practices and rituals.

Narrative sequence

Less is known about how cultural information in prompts affects the same story across scenes.

RESEARCH QUESTION

How does culturally grounded narrative prompting affect cultural representation in AI-generated visual stories?

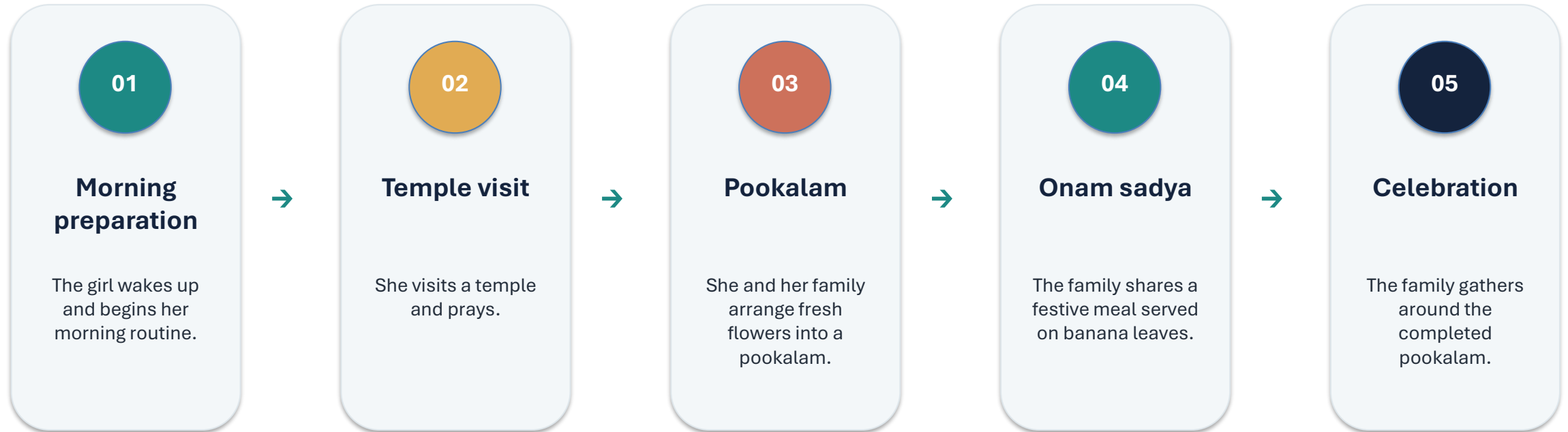
Comparative study design

One five-scene story; two prompt conditions; a controlled generation workflow.



Key controls: same character, scene order, visual style, video model, scene duration, and motion prompts.

The story: “An Onam Morning”



Same character and story progression in both conditions; cultural detail in the image prompts is the key difference.

What changed between the two conditions?

GENERIC NARRATIVE PROMPT

Describes the action without explicit cultural references.

Scene 4 example: a family shares a homemade vegetarian meal using ordinary plates and serving dishes.

CULTURALLY GROUNDED PROMPT

Describes the same action with Onam-specific details.

Scene 4 example: Onam sadya on banana leaves, Kerala clothing and home surroundings, festive food, and eating with hands.

Keyframes

Nano Banana



Video generation

Kling AI



Assembly

5 × 5-second clips



Comparison

Two ~25-second videos

Control: corresponding scenes used the same motion prompt and five-second duration.

How cultural representation was evaluated

01

Cultural fidelity

Accuracy of culturally specific objects, clothing, practices, and activities.

02

Cultural contextuality

Appropriateness of cultural elements within their social and physical setting.

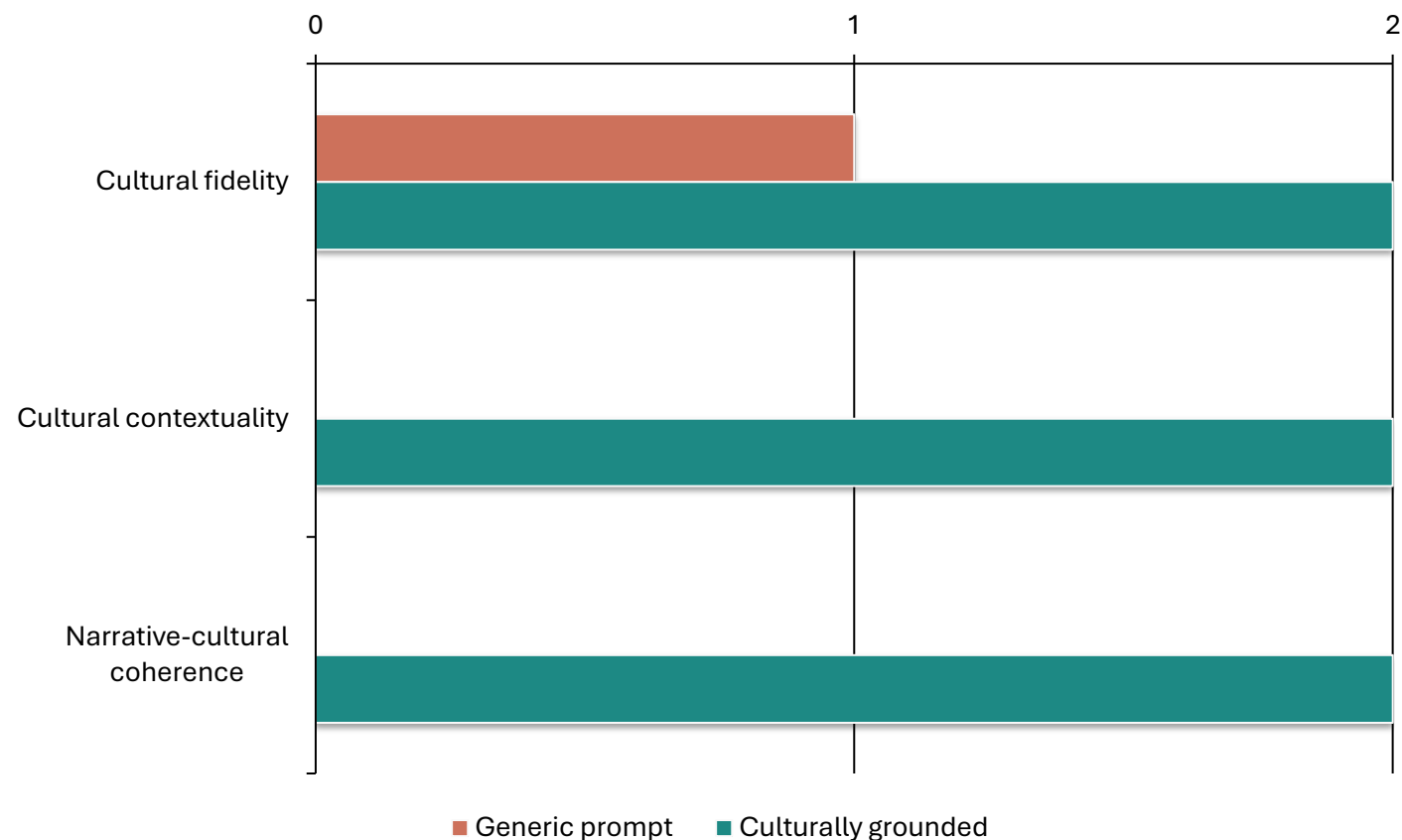
03

Narrative-cultural coherence

Integration of cultural elements with the events and progression of the story.

Scene-level score: 0 = absent/incorrect • 1 = partial/weak • 2 = clear/appropriate/integrated

Results: grounded prompts scored higher on all three dimensions



KEY OBSERVATION

The grounded condition received the score (2/2) in every dimension.

The generic version showed partial cultural fidelity (1), but scored 0 for contextuality and narrative-cultural coherence.

Preliminary support for H1, H2, and H3.

Scores reported in the paper's comparative evaluation; 0–2 scale.

Scene 4 comparison: Onam sadya



The contrast in this scene illustrates why cultural evaluation should consider context and narrative integration—not only whether eating a meal is shown in the image.

Conclusion and next steps

MAIN TAKEAWAY

Adding *explicit cultural information* to narrative image prompts was associated with stronger cultural representation in this comparison.

WHY IT MATTERS

Cultural elements need to be accurate, situated appropriately, and integrated with the story's events.

LIMITATIONS TO ACKNOWLEDGE

- Single story and cultural setting (Onam in Kerala).
- Three dimensions scored on a compact 0–2 scale.
- The paper reports a comparative evaluation, not a large participant study.

NEXT STEPS

- Use multiple cultural narratives and generation runs.
- Include independent raters and report agreement.
- Test whether findings generalise across models and cultures.

Thankyou

To The International Conference on Culture Technology ICCT 2026
And Shanghai University of Engineering Science,China

Thank you