

# Subjective and Objective Descriptions of Driving Scenes in Support of Driver-Automation Interactions

Christopher D. D. Cabrall, Riender Happee, Joost C. F. de Winter  
Delft University of Technology

## Introduction

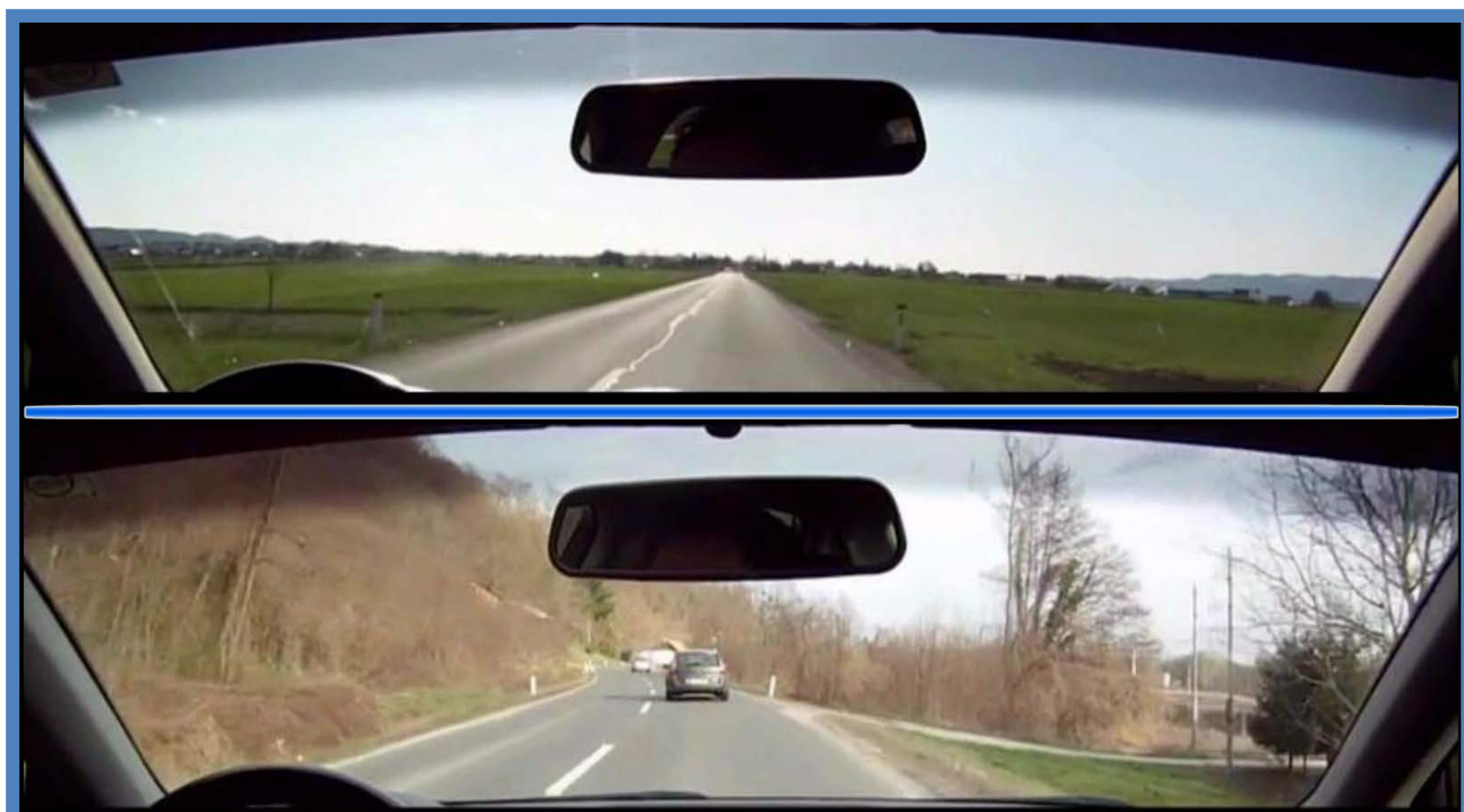
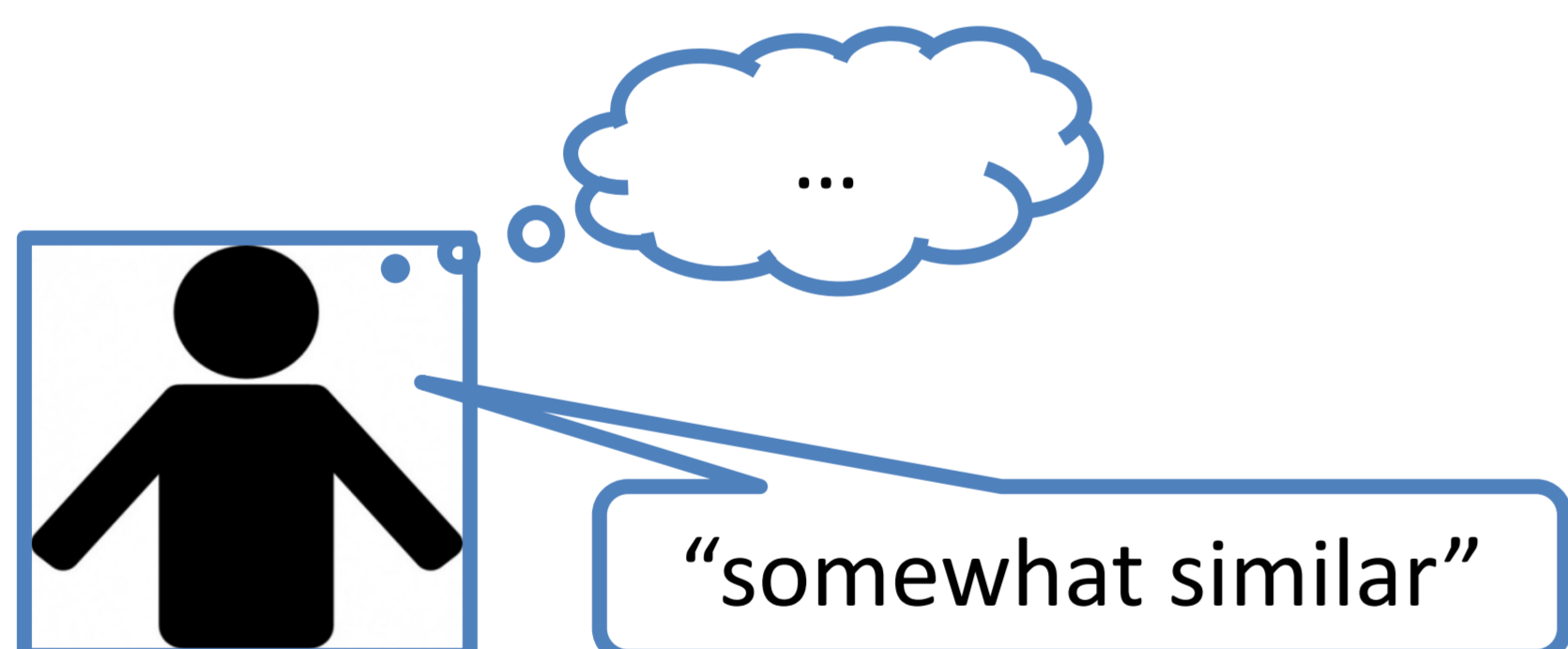
**Background.** Recent advances in the growing domain of automated driving suggest the need for thoughtful design of human-computer interaction strategies. For example, human drivers can process scene variability on implicit levels, but automated systems require explicit rule-based judgments of similarity and difference. What level of abstraction an automation uses in its visual perception may mean the difference between effective human-automation communication, or “uncanny valley”-like conflicts leading to problems of automation disuse, misuse, or abuse.

**Purpose of study.** In the present research, different quantifications (**semantic coding** vs. **computer vision features**) of driving scene-to-scene similarity and difference were compared against **intuitive human judgments** as a reference point for future human-automation interactions.

## Methods

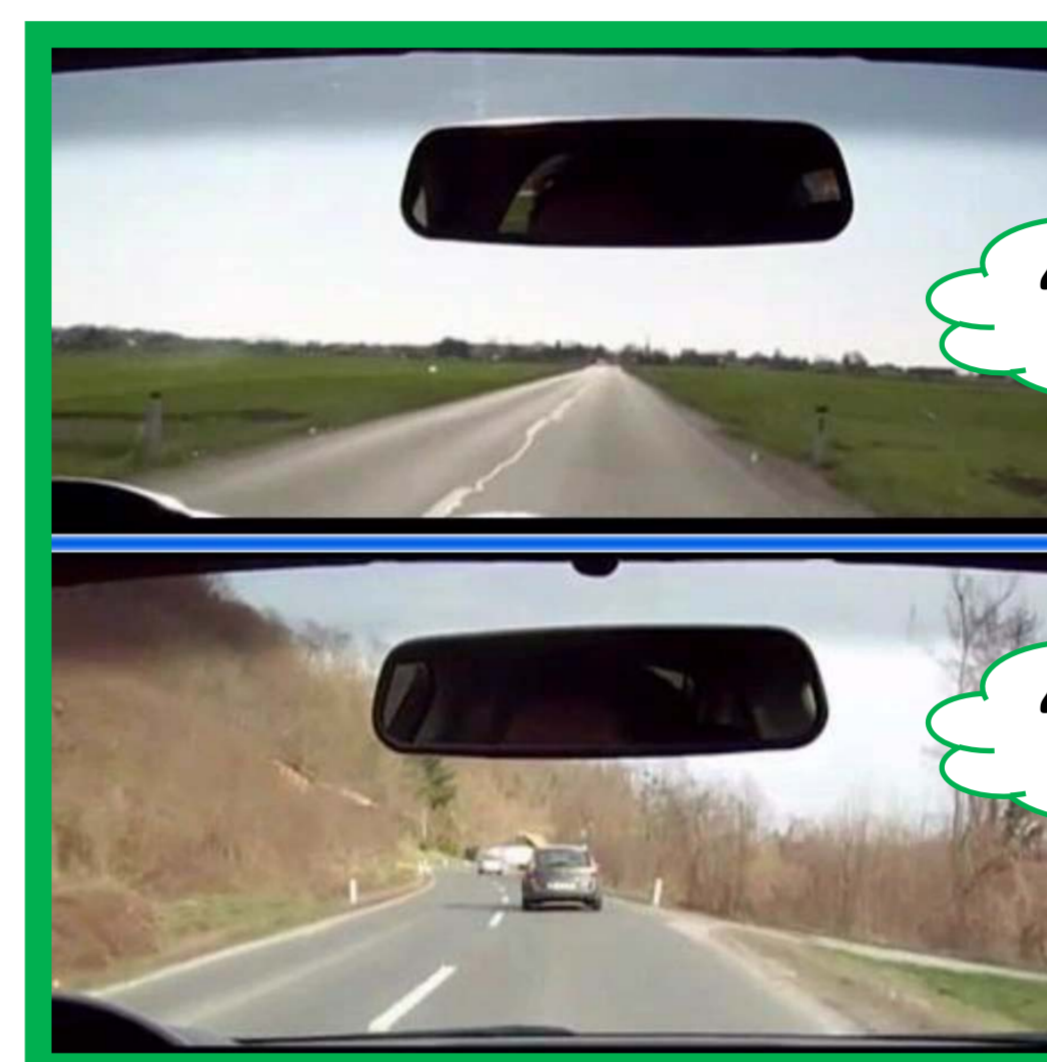
**Participants.** 12 MSc students (11 male : 1 female)  
Mean age = 22.9 yrs old (SD = 1.4)  
Mean driving license = 4.8 yrs (SD = 1.9)

**Procedure.** Each participant rated the same 100 randomly paired driving video clips (i.e., 3 seconds long) on a scale from “0 – Very Different” to “9 – Very Similar”

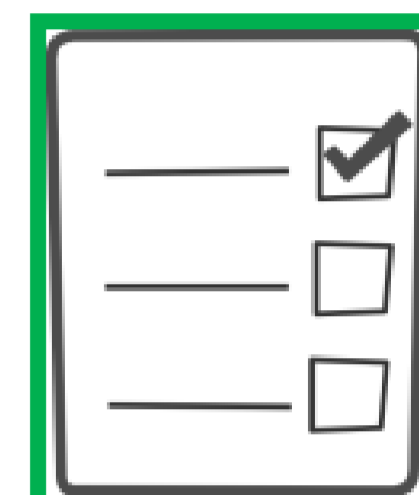


Road curve? (0 or 1)  
Traffic? (0 or 1)  
Misc. details? (0 or 1)  
intersection, stopping,  
lane change, signs, paint

Level 3	111
Level 2	110, 101, 011
Level 1	100, 010, 001
Level 0	000



very similar	0 is within same level
somewhat similar	1 is one level away
somewhat different	2 is two levels away
very different	3 is three levels away



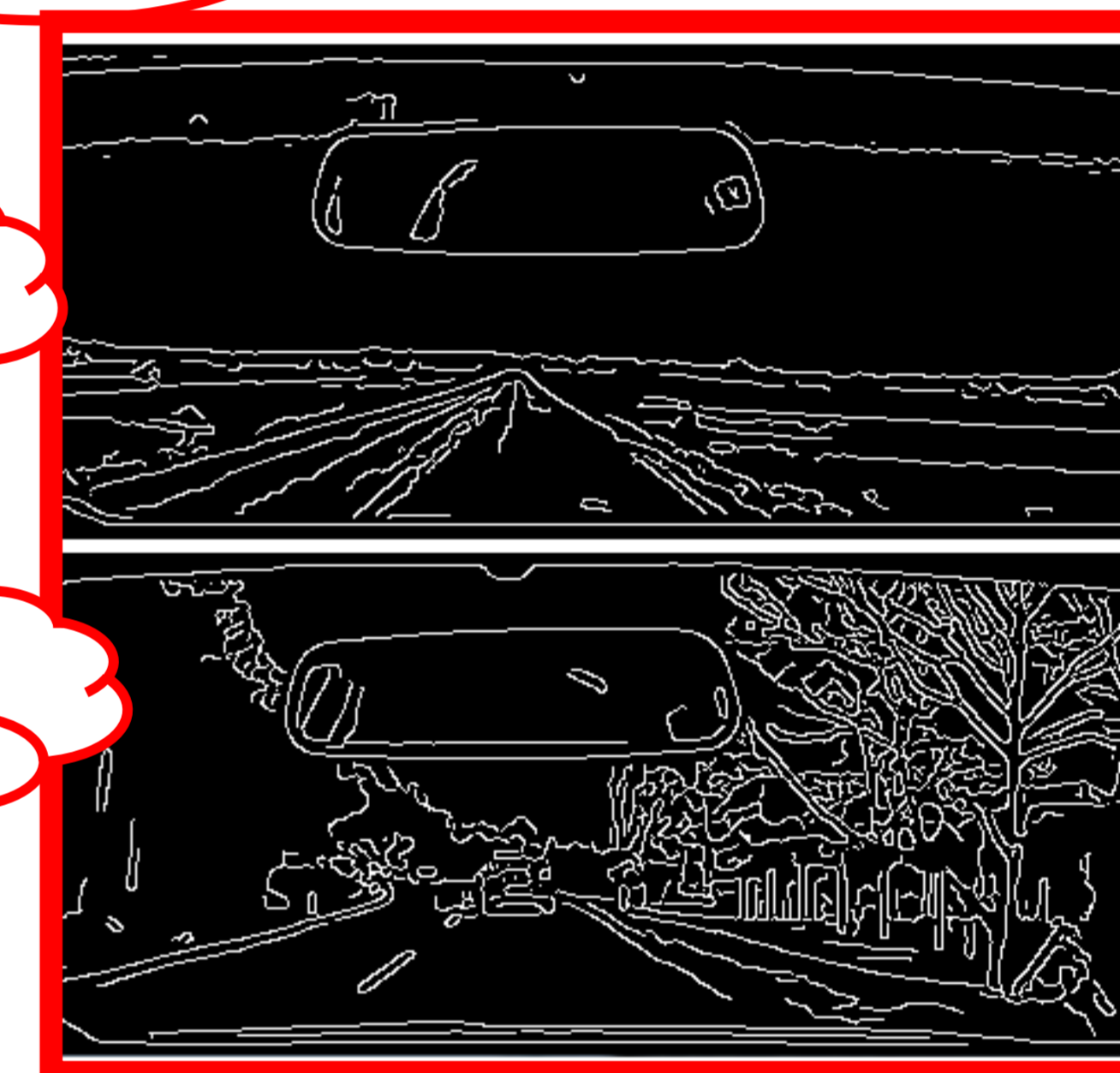
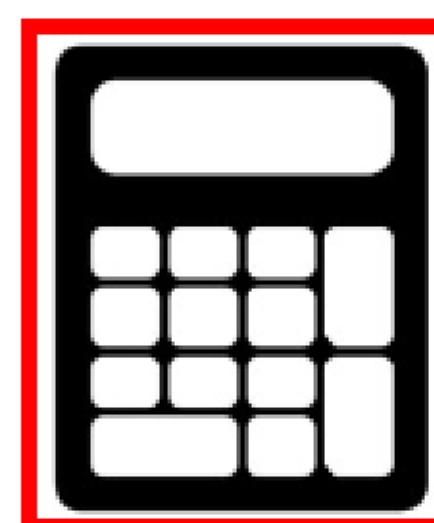
“somewhat different”

Output = edge(myImage,'Canny');

White pixel count?

“9708 pxls”  
= 5.5%

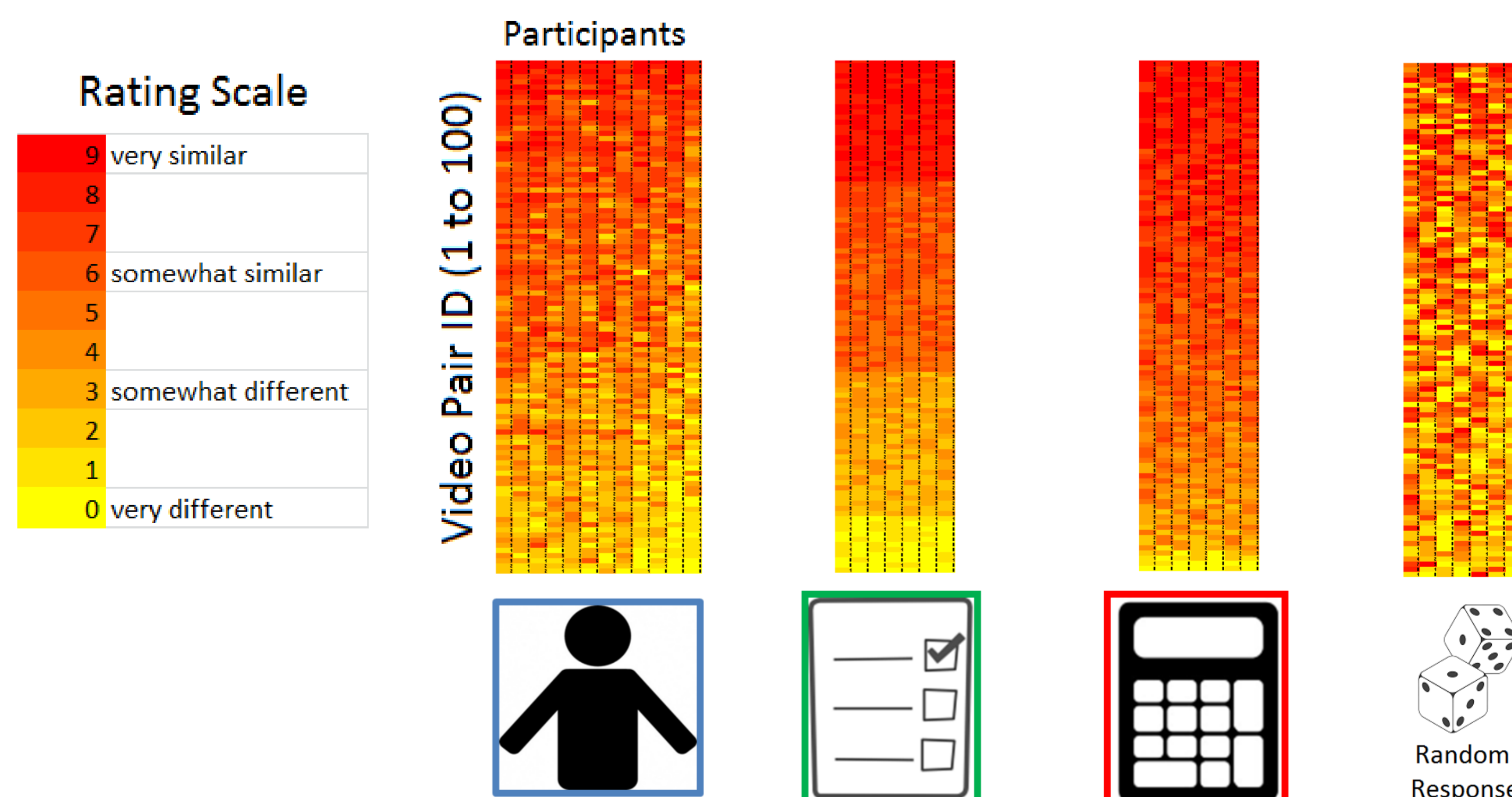
“16635 pxls”  
= 9.4%



“somewhat similar”

## Results/Conclusions

Scene similarity/difference ratings from **semantic coding** quantification showed **closer matches** to **human participant** judgments than those generated from **computer vision**.



Humans evidence apparent non-random individual differences in judging various driving scenes. Both ‘meaning’ and particularly ‘feature’ level descriptions require improvements to coordinate common ground with human intuition of driving scene similarity/difference.