Vision Research 109 (2015) 125–138

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Seeing liquids from visual motion

Takahiro Kawabe [a,*], Kazushi Maruya [a], Roland W. Fleming [b], Shin'ya Nishida [a]

[a] NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, 3-1, Morinosato Wakamiya, Atsugi 243-0198, Kanagawa, Japan
[b] Department of Experimental Psychology, University of Gießen, FB06 – Psychologie, Otto-Behaghel-Str. 10/F2-338, 53594 Gießen, Germany

ARTICLE INFO

Article history:
Received 31 January 2014
Received in revised form 14 July 2014
Available online 4 August 2014

Keywords:
Material perception
Viscosity
Liquid impression
Visual motion

ABSTRACT

Most research on human visual recognition focuses on solid objects, whose identity is defined primarily by shape. In daily life, however, we often encounter materials that have no specific form, including liquids whose shape changes dynamically over time. Here we show that human observers can recognize liquids and their viscosities solely from image motion information. Using a two-dimensional array of noise patches, we presented observers with motion vector fields derived from diverse computer rendered scenes of liquid flow. Our observers perceived liquid-like materials in the noise-based motion fields, and could judge the simulated viscosity with surprising accuracy, given total absence of non-motion information including form. We find that the critical feature for apparent liquid viscosity is local motion speed, whereas for the impression of liquidness, image statistics related to spatial smoothness—including the mean discrete Laplacian of motion vectors—is important. Our results show the brain exploits a wide range of motion statistics to identify non-solid materials.

1. Introduction

Humans can visually recognize not only "things" (objects) that have specific forms, but also "stuff" (materials) that often have no specific form (Adelson, 2001). Over the last decade, the question of how humans perceive materials has received increasing attention (Adelson, 2001; Fleming, 2014; Fleming, Dror, & Adelson, 2003; Fleming, Jäkel, & Maloney, 2011; Kim, Marlow, & Anderson, 2012; Motoyoshi et al., 2007; Nishida & Shinya, 1998; Zaidi, 2011). While previous study of material perception has mainly considered solid materials, many materials around us are in the form of a liquid. For instance, water, the most common liquid, covers 71% of the Earth's surface, and is vital for all known forms of life (CIA, The world fact book). While visual material perception depends on the mechanical properties of a material's body as well as optical properties of a material's surface (Adelson, 2001), past material research has completely ignored how mechanical properties are processed. Liquid viscosity is a critical mechanical property for discriminating water from other liquids. To explore this new direction of material perception research, we examined how human observers recognize liquids and their viscosity from visual information (Kersten, 2011).

One computational scheme for understanding visual processing is physics-based inverse-optics in which relevant physical parameters of visual images are estimated by backtracking the image formation process (Marr, 1975). For many problems of material perception, however, correct inverse computation seems to be impossibly difficult. This is particularly true for the present case, since the movements of liquid particles are not directly observable, and fluid dynamics is chaotically complex. Despite this, in everyday life, we seem to be quite good at perceiving and distinguishing a wide variety of liquids and gels, suggesting the human visual system somehow manages to extract diagnostic information from the retinal images. We reason that complex material properties such as liquidness and viscosity are likely to be estimated from numerous visual cues that *correlate* with the physical properties of interest. The recognition of liquids, and estimation of their properties, presumably draws on mechanisms involved in several visual attributes such as motion, form, and depth, as well as in non-visual attributes including touch sensation. Although these correlations may be imperfect, they may nonetheless provide the brain with a sufficiently reliable source of information to support everyday recognition and interaction. To understand how those cues are combined into a final percept, one should analyse the characteristics of each cue. Here, we focus on the role of motion flow information in liquid and viscosity perception (and report on the role of form information elsewhere; Paulun et al., submitted for publication).

As compared to other visual attributes such as form and color, motion has been believed to play relatively minor roles in "what" processing (visual recognition). Motion information can contribute to recognition of some dynamic objects, but this has been shown only under conditions where object form information is also available (e.g., Chuang, Vuong, & Bülthoff, 2012; Giese & Poggio, 2003;

* Corresponding author. Fax: +81 46 240 4716.
E-mail address: kawabe.takahiro@lab.ntt.co.jp (T. Kawabe).

http://dx.doi.org/10.1016/j.visres.2014.07.003
0042-6989/© 2014 The Authors. Published by Elsevier Ltd.

Johansson, 1973; Lu, 2010; Newell, Wallraven, & Huber, 2004; Stone, 1998; Troje, 2008; Vuong & Tarr, 2006). Here we show that liquid perception is a particularly acute case in which motion analysis can be a central process for visual recognition.

By combining computer graphics, image processing and psychophysics, we have examined whether human observers can perceive liquids and their viscosity solely from optical motion flow. We presented pure motion flow using an array of noise patches, each of which signalled a specific localized motion. This pattern contained no static form information, yet when it reproduced image motion flow corresponding to a dynamic liquid, observers could judge the liquid's viscosity with reasonable accuracy. In this situation, the critical feature for liquid viscosity was found to be local motion speed—the faster the local motion, the less viscous the flow appeared. In addition, spatial smoothness of motion vectors, which can be computationally characterized using the mean discrete Laplacian of motion vectors (see Section 7.2.5 for details), was found to be a critical parameter for creating an impression of a liquid *per se* (i.e., as opposed to some other, non-liquid, source of motion). Our results demonstrate how the visual system exploits image motion statistics for visual understanding of liquids and their viscosity.

2. Experiment 1: liquid viscosity from motion flow

2.1. Purpose

The first experiment examined how well human observers could judge liquid viscosity from dynamic scenes, and from pure motion fields extracted from the same scenes. First, we created 50 computer graphics (CG) movies that simulated ten scenes of opaque liquid flows with five levels of kinematic viscosity for each scene (left panel of Fig. 1A).

Next, we extracted optical flow fields from the simulated movies of liquid motion. Note that the extracted optical flows did not correspond to the physical motion flows of liquid particles—we did not use the flows of liquid particles for visual stimuli, since they were invisible to the observer. We were interested in the image flows visible to the observer. The optical fields were spatially sampled with a 15×15 matrix, and were applied to a two-dimensional array of local noise motion patches, which we called the simulated motion field (right panel in Fig. 1A, see also Movie 2). For each noise patch within the simulated motion field, the carrier (noise) moved at the sampled direction and speed, while the circular envelope remained stationary. Since there was no static form information, this allowed us to isolate the contribution of image motion in estimating liquid viscosity. We asked observers to rate the apparent viscosity both of the original CG movies and of the simulated motion fields with a 5-point scale.

2.2. Methods

Unless otherwise noted, the same methods were used in the subsequent experiments.

2.2.1. Observers

Twenty-eight naive observers (i.e., two groups of fourteen observers) participated in Experiment 1. Half of them participated in the session with CG movies and the other half participated in the session with the simulated motion field. All observers in this study, except the two authors who participated in experiments 3 and 4, were unaware of the specific purpose of experiments. They reported having normal or corrected-to-normal visual acuity. Apart from the authors, participants were paid for the participation. Ethical approval for this study was obtained from the ethical committee at Nippon Telegraph and Telephone Corporation (NTT Communication Science Laboratories Ethical Committee). The experiments were conducted according to the principles laid down in the Helsinki Declaration. Written informed consent was obtained from all participants except the authors.

2.2.2. Apparatus

Stimuli were presented on a 21-in. CRT monitor (GDM-F500R, Sony) with a resolution of 1024×768 pixels and a refresh rate of 60 Hz. We linearized the luminance emitted from the monitor in a range from 0 to 132 cd/m^2 using a photometer (OP200-E, Cambridge Research Systems). A computer (Mac pro, Apple) controlled stimulus presentation and data collection with MATLAB and its extension (PsychToolBox 3, Brainard, 1997; Pelli, 1997).

2.2.3. Stimuli

2.2.3.1. Simulation of fluid dynamics. We used a physics engine implemented in Blender (http://www.blender.org/) to simulate fluid dynamics. The resolution of simulation mesh (i.e., the granularity at which the actual fluid simulation is performed; see http://wiki.blender.org/index.php/Doc:2.4/Manual/Physics/Fluid for details) was set to 150. We set gravity to -9.81 m/s^2 along the *Z*-axis. The degree of liquid viscosity was manipulated by changing the kinematic viscosity in five levels (10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, and 10^{0} m^2/s). The surfaces of the simulated liquids were a gray Lambertian material with specular highlights (all three RGB channels were set to 0.8 and surface diffuse reflectivity was set to 0.5) that were calculated using the Cook–Torrance model with intensity of 1.0 and hardness of 150. These surfaces were lit using environmental lighting (Energy: 0.400 with sky color), which is predefined in Blender as a kind of global illumination; we also applied ambient occlusion (Factor: 1.00) and indirect lighting (Factor: 6.00). Based on the simulation, we created movies 2 s long (i.e. 33.3 ms $\times$ 60 frames) of the ten different scenes. The movies subtended 13.6×13.6 deg of visual angle (i.e. 384×384 pixels in the display). In total, ten different scenes were created (see Movie 1). In scene 1, fluids were emitted downward from two sources with a 10 cm $\times$ 10 cm $\times$ 10 cm cubic shape that were located 20 cm left and right of, and 25 cm above, the center of the floor of a 50 cm (*x* dimension) $\times$ 50 cm (*y* dimension) $\times$ 100 cm (*z* dimension) container. The inflow velocity was 5 m/s. The location of the camera viewing the scene was *X*: 0 cm, *Y*: 0 cm, and *Z*: 55 cm, and the rotation of the camera was *X*: 0°, *Y*: 0° and *Z*: 0°. The camera had a focal length of 35 mm (the focal length is constant across scenes). Scene 2 consisted of an open vessel initially containing 34 cm $\times$ 34 cm $\times$ 16 cm of fluid, which tilted by 90 deg causing the fluid to cascade onto the floor of another container of 100 cm $\times$ 100 cm $\times$ 60 cm. The location of the camera was *X*: 0 cm, *Y*: -80 cm, and *Z*: 55 cm, and the rotation of the camera was *X*: 60°, *Y*: 0° and *Z*: 0°. In scene 3, fluid was emitted from a 100 cm $\times$ 100 cm $\times$ 4 cm source onto a 45° slanted floor, and flowed along the floor surface naturally. The inflow velocity was 4 m/s. The location of the camera was *X*: 100 cm, *Y*: 0 cm, and *Z*: 50 cm, and the rotation of the camera was *X*: 90°, *Y*: 0° and *Z*: 90°. In scene 4, a sphere of fluid with a 40 cm diameter was thrown toward the wall of a 100 cm $\times$ 100 cm $\times$ 100 cm container. The location of the camera was *X*: 0 cm, *Y*: 80 cm, and *Z*: 100 cm, and the rotation of the camera was *X*: 55°, *Y*: 0° and *Z*: 180°. In scene 5, two sources with a 10 cm $\times$ 10 cm $\times$ 10 cm cubic shape moved inside a 50 cm $\times$ 50 cm $\times$ 100 cm container, and emit fluids onto the floor of the container. The inflow velocity was 5 m/s. The location of the camera was *X*: 0 cm, *Y*: -25 cm, and *Z*: 20 cm, and the rotation of the camera was *X*: 50°, *Y*: 0° and *Z*: 0°. In scene 6, fluid was emitted from a source with a 10 cm $\times$ 10 cm $\times$ 10 cm cubic shape toward the wall of a 100 cm $\times$ 100 cm $\times$ 100 cm container. The location of the camera was *X*: 50 cm, *Y*: 50 cm, and *Z*: 20 cm, and the